

```

XX PN US5869291-A.
XX PD 09-FEB-1999.
XX PF 27-DEC-1996; 96US-0773423.
XX PR 27-DEC-1996; 96US-0773423.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Bandman O, Goli SK, Hillman JL;
XX DR WPI; 1999-152769/13.
XX DR N-PSDB; AAX04755.
XX PT DNA encoding human Rab proteins - useful for producing recombinant
XX PT proteins
XX PS Example 1; Fig 3A-B; 36pp; English.
XX CC The present sequence represents Rab protein designated HRAB-3. The
XX CC nucleic acid sequence was first identified in Incyte clone 641412
XX CC from a breast tissue cDNA library. The HRAB nucleic acid sequences
XX CC are useful for producing recombinant human Rab proteins which may
XX CC be useful for treating disorders involving vesicle targeting,
XX CC membrane transfer or fusion, or protein processing, targeting or
XX CC secretion, e.g. disorders characterised by abnormal neurotransmitter
XX CC production and release, lysosomal storage diseases, chorioidaemia and
XX CC endocrine hypofunction and hyperfunction.
XX SQ Sequence 201 AA;
Query Match 100.0%; Score 1062; DB 20; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.8e-106;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGKSLFKVILLGGVGKSLMRYNTKPDFTQFTIGVEFLNKDLEVDGHFTVQI 60
DB 1 MAGKSLFKVILLGGVGKSLMRYNTKPDFTQFTIGVEFLNKDLEVDGHFTVQI 60
QY 61 WDTAGQERFRLTPFYRGSDCLLTFSVDDSQSFQNLNWKKEFIYYADYKEPSPFFV 120
DB 61 WDTAGQERFRLTPFYRGSDCLLTFSVDDSQSFQNLNWKKEFIYYADYKEPSPFFV 120
QY 121 ILGNKIDISERQVSTEEAQAACRNDGYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
DB 121 ILGNKIDISERQVSTEEAQAACRNDGYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSSCC 201
DB 181 HLIQTDTVNLHRKPKPSSSCC 201
RESULT 2
ID ABP53078 standard; Protein; 201 AA.
XX AC ABP53078;
XX DT 08-NOV-2002 (first entry)
XX DE Human Rab protein HRAB-3 SEQ ID NO:5.
XX KW Human; Rab protein; HRAB-3; antianaemic; antidiabetic; antiinflammatory;
XX KW immunosuppressive; dermatological; neuroprotective; antiulcer; anti-HIV;
XX KW antirheumatoid; antiarthritic; cyostatic; endocrine; gene therapy;
XX KW immune system disorder; autoimmune haemolytic anaemia; myasthenia gravis;
XX KW juvenile diabetes mellitus; lupus erythematosus; multiple sclerosis;
XX KW Crohn's disease; rheumatoid arthritis; AIDS; ulcerative colitis; cancer;
XX KW leukaemia; lymphoma; carcinoma; vesicle targeting disease; ras family;
XX KW lysosomal storage disease; chorioidaemia; endocrine hypofunction;
XX KW endocrine hyperfunction.

```

```

XX OS Homo sapiens.
XX PN US2002090712-A1.
XX PD 11-JUL-2002.
XX PF 19-NOV-2001; 2001US-0988974.
XX PR 27-DEC-1996; 96US-0773423.
XX PR 17-DEC-1998; 98US-0215887.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Hillman JL, Goli SK, Bandman O;
XX DR WPI; 2002-655836/70.
XX DR N-PSDB; ABQ75524.
XX PT Novel Rab polypeptides, low molecular weight guanidine triphosphatases
XX PT of Ras superfamily, useful for treating, preventing or diagnosing
XX PT cancer, and immune system disorders e.g. ulcer, arthritis and Crohn's
XX PT disease.
XX PS Claim 1; Fig 3A-B; 40pp; English.
XX CC The present sequence represents human HRAB-3 which is a Rab protein (low
XX CC molecular weight guanidine triphosphatases of the Ras superfamily).
XX CC Rab proteins (I) have antianaemic, antidiabetic, immunosuppressive,
XX CC antiinflammatory, dermatological, neuroprotective, antirheumatoid,
XX CC antiulcer, antiarthritic, anti-HIV, cyostatic and endocrine activities,
XX CC and can be used in gene therapy. (I) can be used for screening a compound
XX CC for effectiveness as an agonist or antagonist, for screening a compound
XX CC that specifically binds to (I) or modulates the activity of (I), and for
XX CC preparing a polyclonal or monoclonal antibody by hybridoma technology.
XX CC (I) can be used for diagnosing, treating or preventing immune system
XX CC disorders (including autoimmune haemolytic anaemia, juvenile diabetes
XX CC mellitus, lupus erythematosus, multiple sclerosis, myasthenia gravis,
XX CC Crohn's disease, rheumatoid arthritis, AIDS or ulcerative colitis),
XX CC cancer (including leukaemia, lymphoma, carcinomas of lung, penis, breast,
XX CC pancreas, colon, stomach, small intestine, brain or prostate), diseases
XX CC involving vesicle targeting, membrane transfer or fusion, or protein
XX CC processing, targeting or secretion, and conditions characterised by
XX CC abnormal levels of neurotransmitter production and release, lysosomal
XX CC storage diseases, chorioidaemia, endocrine hypofunction and
XX CC hyperfunction.
XX SQ Sequence 201 AA;
Query Match 100.0%; Score 1062; DB 23; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.8e-106;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAGKSLFKVILLGGVGKSLMRYNTKPDFTQFTIGVEFLNKDLEVDGHFTVQI 60
DB 1 MAGKSLFKVILLGGVGKSLMRYNTKPDFTQFTIGVEFLNKDLEVDGHFTVQI 60
QY 61 WDTAGQERFRLTPFYRGSDCLLTFSVDDSQSFQNLNWKKEFIYYADYKEPSPFFV 120
DB 61 WDTAGQERFRLTPFYRGSDCLLTFSVDDSQSFQNLNWKKEFIYYADYKEPSPFFV 120
QY 121 ILGNKIDISERQVSTEEAQAACRNDGYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
DB 121 ILGNKIDISERQVSTEEAQAACRNDGYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSSCC 201
DB 181 HLIQTDTVNLHRKPKPSSSCC 201
RESULT 3
ABP31125
ID ABB31125 standard; Peptide; 209 AA.

```

XX ABB31125;
 XX AC 01-FEB-2002 (first entry)
 XX DT
 XX DE Peptide #3776 encoded by breast cell single exon nucleic acid probe.
 XX KW Human; microarray; single exon probe; gene expression; breast;
 XX KW disease; cancer.
 XX OS Homo sapiens.
 XX PN WO200157271-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00662.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-496933/54.
 XX XX New spatially-addressable set of single exon nucleic acid probes,
 PT useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 XX
 PS Claim 27; SEQ ID NO 14093; 327pp + sequence listing; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a peptide encoded by a single exon nucleic acid
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 209 AA;
 Query Match 100.0%; Score 1062; DB 22; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGKSLFKVILLGGGVGKSLMRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFTVMTQI 60
 DB 9 MAGKSLFKVILLGGGVGKSLMRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFTVMTQI 68
 QY 61 WDTAGQERFRSLRTPFYRGSDCLLTFSVDDSQSFQNLNNWKEFIYYADVKEPESFPFV 120
 DB 69 WDTAGQERFRSLRTPFYRGSDCLLTFSVDDSQSFQNLNNWKEFIYYADVKEPESFPFV 128
 QY 121 ILGNKIDISERQVSTEEAQWCRNDGYPYFETSAKDNTNVAAPFEAVRVLATEDRSD 180

DB 129 ILGNKIDISERQVSTEEAQWCRNDGYPYFETSAKDNTNVAAPFEAVRVLATEDRSD 188
 QY 181 HLIQTDVNLHRKPKPSSSCC 201
 DB 189 HLIQTDVNLHRKPKPSSSCC 209
 RESULT 4
 AAM57088
 ID AAM57088 standard; Protein; 209 AA.
 XX AC AAM57088;
 XX DT 05-NOV-2001 (first entry)
 XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 29193.
 XX KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer.
 OS Homo sapiens.
 XX PN WO200157275-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00667.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-493446/52.
 XX XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PS Example 4; SEQ ID NO: 29193; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 209 AA;
 Query Match 100.0%; Score 1062; DB 22; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGKSLFKVILLGGGVGKSLMRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFTVMTQI 60
 DB 9 MAGKSLFKVILLGGGVGKSLMRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFTVMTQI 68
 QY 61 WDTAGQERFRSLRTPFYRGSDCLLTFSVDDSQSFQNLNNWKEFIYYADVKEPESFPFV 120
 DB 69 WDTAGQERFRSLRTPFYRGSDCLLTFSVDDSQSFQNLNNWKEFIYYADVKEPESFPFV 128
 QY 121 ILGNKIDISERQVSTEEAQWCRNDGYPYFETSAKDNTNVAAPFEAVRVLATEDRSD 180

Db 129 ILGNKIDISERQVSTTEAQCWCRNGDYPYFETSAKDATNVAAPFEAVRVLATEDRSD 188
 QY 181 HLIQTDVTNLRKPKPSSCC 201
 Db 189 HLIQTDVTNLRKPKPSSCC 209

RESULT 5
 AAM17322
 ID AAM17322 standard; Protein; 209 AA.
 AC AAM17322;
 DT 12-OCT-2001 (first entry)
 DE Peptide #3756 encoded by probe for measuring cervical gene expression.
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX Homo sapiens.
 OS
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID No 22148; 487pp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP: see AAI10068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 209 AA;
 Query Match 100.0%; Score 1062; DB 22; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGGVGKSLMNRVYTNKFDTLQFHTIGVEFLNKDLVDGHHVTMQI 60
 Db 9 MAGKSLFKVILLGGVGKSLMNRVYTNKFDTLQFHTIGVEFLNKDLVDGHHVTMQI 68

QY 61 WDTAGQERFRSLRTPPYGSDCLLTFVSVDQSQFNLNWKKEFIYYADYKESPEFPV 120
 Db 69 WDTAGQERFRSLRTPPYGSDCLLTFVSVDQSQFNLNWKKEFIYYADYKESPEFPV 128

QY 121 ILGNKIDISERQVSTTEAQCWCRNGDYPYFETSAKDATNVAAPFEAVRVLATEDRSD 180
 Db 129 ILGNKIDISERQVSTTEAQCWCRNGDYPYFETSAKDATNVAAPFEAVRVLATEDRSD 188

QY 181 HLIQTDVTNLRKPKPSSCC 201
 Db 189 HLIQTDVTNLRKPKPSSCC 209

RESULT 6
 AAM05002
 ID AAM05002 standard; Protein; 209 AA.
 AC AAM05002;
 DT 09-OCT-2001 (first entry)
 DE Peptide #3684 encoded by probe for measuring breast gene expression.
 KW Probe; human; breast disease; breast cancer; development disorder;
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
 XX Homo sapiens.
 OS
 PN WO200157270-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 29-JAN-2001; 2001WO-US00661.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-476286/51.
 DR
 XX Novel single exon nucleic acid probe used to measuring gene expression
 PT in a human breast -
 XX
 PS Claim 27; SEQ ID No 13742; 322pp; English.
 XX
 CC The present invention relates to novel single exon nucleic acid probes
 CC (see AAI00010-AA110067). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for measuring human gene expression in
 CC a human breast sample, where the probe hybridises at high stringency to a
 CC nucleic acid expressed in the human breast. The probes are useful for
 CC predicting, diagnosing, grading, staging, monitoring and prognosis
 CC diseases of the human breast, particularly those diseases with polygenic
 CC aetiology. The diseases include: breast cancer, disorders of development,
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative
 CC breast disease and non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 209 AA;
 Query Match 100.0%; Score 1062; DB 22; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGGVGKSLMNRVYTNKFDTLQFHTIGVEFLNKDLVDGHHVTMQI 60
 Db 9 MAGKSLFKVILLGGVGKSLMNRVYTNKFDTLQFHTIGVEFLNKDLVDGHHVTMQI 68

QY 61 WDTAGQERFSLRTPFYRGSDCLLTFVSDDQSQFNLNNKKEFIYYADVKEPESFPFV 120
 DB 69 WDTAGQERFSLRTPFYRGSDCLLTFVSDDQSQFNLNNKKEFIYYADVKEPESFPFV 128
 QY 121 ILGNKIDISERQVSTEEAQAQCRDNGDYPYFETSADKATNVAAPFEAVRRVLATERDSD 180
 DB 129 ILGNKIDISERQVSTEEAQAQCRDNGDYPYFETSADKATNVAAPFEAVRRVLATERDSD 188
 QY 181 HLIQDTVNLHRKPKPSSCC 201
 DB 189 HLIQDTVNLHRKPKPSSCC 209

RESULT 7
 ID ABG39111 standard; Peptide; 209 AA.
 AC ABG39111;
 DT 19-AUG-2002 (first entry)
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 28776.
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 XX WO200186003-A2.
 XX 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX Claim 27; SEQ ID No 28776; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the labeled detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryotic lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 209 AA;
 SQ

Query Match 100.0%; Score 1062; DB 23; Length 209;
 Best Local Similarity 100.0%; Pred. No. 4e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGDGVGKSSLMNRYVTKKPTQLFHTIGVEFLNKDLEVDGHFTVMOI 60
 DB 9 MAGKSLFKVILLGDGVGKSSLMNRYVTKKPTQLFHTIGVEFLNKDLEVDGHFTVMOI 68
 QY 61 WDTAGQERFSLRTPFYRGSDCLLTFVSDDQSQFNLNNKKEFIYYADVKEPESFPFV 120
 DB 69 WDTAGQERFSLRTPFYRGSDCLLTFVSDDQSQFNLNNKKEFIYYADVKEPESFPFV 128
 QY 121 ILGNKIDISERQVSTEEAQAQCRDNGDYPYFETSADKATNVAAPFEAVRRVLATERDSD 180
 DB 129 ILGNKIDISERQVSTEEAQAQCRDNGDYPYFETSADKATNVAAPFEAVRRVLATERDSD 188
 QY 181 HLIQDTVNLHRKPKPSSCC 201
 DB 189 HLIQDTVNLHRKPKPSSCC 209

RESULT 8
 ABP53081
 ID ABP53081 standard; Protein; 201 AA.
 XX
 AC ABP53081;
 XX
 DT 08-NOV-2002 (first entry)
 XX Canine Rab98 protein SEQ ID NO:9.
 DE
 XX Human; Rab protein; HRAS; antianemic; antidiabetic; antiinflammatory;
 XX immunosuppressive; dermatological; neuroprotective; antiulcer; anti-HIV;
 KW antirheumatoid; antirheumatic; cyostatic; endocrine; gene therapy;
 KW immune system disorder; autoimmune haemolytic anaemia; myasthenia gravis;
 KW juvenile diabetes mellitus; lupus erythematosus; multiple sclerosis;
 KW Crohn's disease; rheumatoid arthritis; AIDS; ulcerative colitis; cancer;
 KW leukaemia; lymphoma; carcinoma; vesicle targeting disease; ras family;
 KW lysosomal storage disease; choroideaemia; endocrine hypofunction;
 XX endocrine hyperfunction.
 XX Canis sp.
 XX

PN US2002090712-A1.
 XX 11-JUL-2002.
 PD 19-NOV-2001; 2001US-0988974.
 XX 27-DEC-1996; 96US-0773423.
 PR 17-DEC-1998; 98US-0215887.
 XX (INCY-) INCYTE PHARM INC.
 PA Hillman JL, Goli SK, Bandman O;
 PI WPI; 2002-655836/70.
 XX Novel Rab polypeptides, low molecular weight guanidine triphosphatases
 PT of Ras superfamily, useful for treating, preventing or diagnosing
 PT cancer, and immune system disorders e.g. ulcer, arthritis and Crohn's
 PT disease -
 XX Disclosure; Fig 4; 40pp; English.
 PS The present invention describes human Rab proteins (low molecular weight
 CC guanidine triphosphatases of the Ras superfamily) designated HRABS.
 CC Rab proteins (I) have antidiabetic, antidiabetic, immunosuppressive,
 CC antiinflammatory, dermatological, neuroprotective, antirheumatoid,
 CC antiulcer, antiarthritic, anti-HIV, cytostatic and endocrine activities,
 CC and can be used in gene therapy. (I) can be used for screening a compound
 CC for effectiveness as an agonist or antagonist, for screening a compound
 CC that specifically binds to (I) or modulates the activity of (I), and for
 CC preparing a polyclonal or monoclonal antibody by hybridoma technology.
 CC (I) can be used for diagnosing, treating or preventing immune system
 CC disorders (including autoimmune haemolytic anaemia, juvenile diabetes
 CC mellitus, lupus erythematosus, multiple sclerosis, myasthenia gravis,
 CC Crohn's disease, rheumatoid arthritis, AIDS or ulcerative colitis),
 CC cancer (including leukaemia, lymphoma, carcinomas of lung, penis, breast,
 CC pancreas, colon, stomach, small intestine, brain or prostate) diseases
 CC involving vesicle targeting, membrane transfer or fusion, or protein
 CC processing, targeting or secretion, and conditions characterised by
 CC abnormal levels of neurotransmitter production and release, lysosomal
 CC storage diseases, chorioideraemia, endocrine hypofunction and
 CC hyperfunction. The present sequence represents a Rab protein given in
 CC comparison with the human HRABS (HRABS-1, HRABS-2 and HRABS-3) from the
 CC present invention.
 XX SQ Sequence 201 AA;
 Query Match 99.5%; Score 1057; DB 23; Length 201;
 Best Local Similarity 99.5%; Pred. No. 1.3e-105;
 Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGKSSLEKVLGGGVGKSSLMNRYVYTNKFDLTQFTHTIGVEFLNKDLEVDGHFTVMOI 60
 DB 1 MAGKSSLEKVLGGGVGKSSLMNRYVYTNKFDLTQFTHTIGVEFLNKDLEVDGHFTVMOI 60
 QY 61 WDTAQERFSLRTPFFYRGSDCCLLTFSVDDSQSQFNLNWKKEFIYYADVKEPESPPFV 120
 DB 61 WDTAQERFSLRTPFFYRGSDCCLLTFSVDDSQSQFNLNWKKEFIYYADVKEPESPPFV 120
 QY 121 ILGNKIDISERQVSTEEAQAACRONGDYPYPTSAKDATNVAAAFEEAVRVLATEDRSD 180
 DB 121 ILGNKIDISERQVSTEEAQAACRONGDYPYPTSAKDATNVAAAFEEAVRVLATEDRSD 180
 QY 181 HLIQDTVNLHRKPKPSSSCC 201
 DB 181 HLIQDTVNLHRKPKPSSSCC 201
 RESULT 9
 ID AAB19172 standard; Protein; 201 AA.
 XX AAB19172
 AC AAB19172;

XX 19-FEB-2001 (first entry)
 DT Amino acid sequence of human RAB34.
 XX RAB protein; GTPase; GTP binding; gene therapy; cancer.
 XX Homo sapiens.
 XX WO200058464-A2.
 XX 05-OCT-2000.
 XX 13-MAR-2000; 2000WO-US06330.
 XX 25-MAR-1999; 99US-0126083.
 XX (AXYS-) AXYS PHARM INC.
 XX Allen M, Abel K, McIntosh B, Vega R, Rutter M, Buckler AJ;
 WPI; 2000-647233/62.
 DR N-PSDB; AA96894.
 XX Novel isolated nucleic acid encoding a mammalian RAB protein useful for
 PT identifying homologous or related genes, in producing composition that
 PT modulates expression or function of RAB for cancer therapy -
 XX Claim 2; Page 50-51; 58pp; English.
 CC The present sequence represents a mammalian RAB protein. RAB proteins
 CC constitute the largest family of small GTPases, with over 40 identified
 CC isoforms. RAB proteins contain four highly conserved peptide sequences
 CC involved in GTP binding and hydrolysis. Compositions comprising RAB
 CC nucleic acid are useful for identifying homologous or related genes,
 CC in producing compositions that modulate the expression or function of
 CC RAB for gene therapy, mapping functional regions of the protein and
 CC in studying associated physiological pathways. In addition, modulation
 CC of the gene activity in vivo is used for prophylactic and therapeutic
 CC purposes, such as treatment of cancer, and identification of cell type
 CC based on expression. The DNA may also be used to identify expression of
 CC the gene in a biological specimen.
 XX SQ Sequence 201 AA;
 Query Match 79.0%; Score 839; DB 21; Length 201;
 Best Local Similarity 76.1%; Pred. No. 4.3e-82;
 Matches 153; Conservative 22; Mismatches 26; Indels 0; Gaps 0;
 QY 1 MAGKSSLEKVLGGGVGKSSLMNRYVYTNKFDLTQFTHTIGVEFLNKDLEVDGHFTVMOI 60
 DB 1 MSGKSLLLKVLGGGVGKSSLMNRYVYTNKFDLSQAFTIGVEFLNRDLVDGFRVLIQI 60
 QY 61 WDTAQERFSLRTPFFYRGSDCCLLTFSVDDSQSQFNLNWKKEFIYYADVKEPESPPFV 120
 DB 61 WDTAQERFSLRTPFFYRGADCCLLTFSVDDSQSQFNLNWKKEFIYYADVKEPESPPFV 120
 QY 121 ILGNKIDISERQVSTEEAQAACRONGDYPYPTSAKDATNVAAAFEEAVRVLATEDRSD 180
 DB 121 VLGNKVDKEDRQVTTTEEAQTWCWENGDPYLETSAKDTNTVTFEAVRQVLAEEGLE 180
 QY 181 HLIQDTVNLHRKPKPSSSCC 201
 DB 181 HCVLGHITDLNSGKAGSSCC 201
 RESULT 10
 ID ABP43693 standard; Protein; 201 AA.
 XX ABP43693
 AC ABP43693;
 XX 26-FEB-2003 (first entry)
 DT

XX RAB9-like protein.
DE
XX Neuroprotective; immunomodulator; cancer; chromosome X;
KW cytosolic; anti-inflammatory; gene therapy; nutritional supplement;
KW wound; burn; ulcer; Alzheimer's disease; Huntington's disease;
KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
KW vulnery.
XX
XX Homo sapiens.
XX WO200231111-A2.
XX 18-APR-2002.
XX 11-OCT-2001; 2001WO-US27760.
XX 12-OCT-2000; 2000US-0687527.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-426278/45.
XX N-PSDB; ABQ60937.
XX
XX New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation
XX
XX Claim 20; SEQ ID # 596; 357pp + sequence listing; English.
XX The invention relates to 446 newly isolated polynucleotide sequences.
CC The activity of polynucleotides of the invention may be described as,
CC vulnery, neuroprotective, immunomodulator, cytostatic and
CC anti-inflammatory. Compositions comprising nucleic acids of the invention
CC are useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records
CC AB43544-AB43989 represent polypeptides encoded by polynucleotides of
CC the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ, Sequence 201 AA;
Query Match 79.0%; Score 839; DB 23; Length 201;
Best Local Similarity 76.1%; Pred. No. 4.3e-82;
Matches 153; Conservative 22; Mismatches 26; Indels 0; Gaps 0;
QY 1 MAGKSSLEKVLGGGKSLMNRYYTNKFDQFHTIGVEFLNKDLEVDGHEFTMOI 60
Db 1 MSGKSELLKVLGGGKSLMNRYYTNKFDQFHTIGVEFLNKDLEVDGHEFTMOI 60
QY 61 WDTAQERFSRLTPFFYRGSDCLLITSDVDSQFNLNWKKEFIYYADVKEPSPFV 120
Db 61 WDTAQERFSRLTPFFYRGSDCLLITSDVDSQFNLNWKKEFIYYADVKEPSPFV 120
QY 121 ILGNKIDISERQVSTEEAQACRNGDYPYFETSAKATNVAFAFBAVRVLATEPDS 180
Db 121 VLGNKVDKEDQVSTEEAQVCMENGDPYLETSAKODTNTVFAFBAVRVLAVEQLE 180
QY 181 HLIQDTVNLNRKPKPSSCC 201
Db 181 HCMLGHTIDLNSGKAGSSCC 201

RESULT 11
AAU17556
ID AAU17556 standard; Protein; 222 AA.
XX
XX AAU17556;
AC
DT 07-NOV-2001 (first entry)
XX
DE Novel signal transduction pathway protein, Seq ID 1121.
XX
XX Neuroprotective; cytostatic; dermatological; immunosuppressive; tumour;
KW antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; inflammatory condition;
KW organ transplant rejection; infection; hepatitis C; blood disorder;
KW sickle cell anaemia; hyperproliferative disorder; Gaucher's disease;
KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS;
KW acquired immune deficiency syndrome.
XX
XX Homo sapiens.
XX WO200154733-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01312.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 08-SEP-2000; 2000US-0231968.
PR 12-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235835.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 20-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246533.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.

PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-465460/50.
DR N-PSDB; AAS27473.
XX Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX Claim 1; SEQ ID No 1121; 880pp; English.
XX The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC disorders (e.g. Alzheimer's disease, Parkinson's disease), chromosomal
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders (e.g.
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Query Match 79.0%; Score 839; DB 22; Length 222;

Best Local Similarity 76.1%; Pred. No. 4.9e-82;
Matches 153; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGGGVGKSSLMRYNTNKPDTLFTTIGVEFLNKDLVDGHEVTWQI 60
Db 22 MSGKSILLKVVILLGGGVGKSSLMRYNTNKPDSQAFHTIGVEFLNRDLVDGHEVTQI 81
QY 61 WDTAGQERPSLRTPFYRGSDCLLTFSVDDSSQSNLNNKKEFIYADVKPEPSFPV 120
Db 82 WDTAGQERPSLRTPFYRGADCLLTFSVDDRSQSNLNNKKEFIYADVKDPHFPV 141
QY 121 ILGNKIDISEROVSTEEAQACRDNGDYPYPTSAKATNVAAAFEEAVRRVLATEDRSD 180
Db 142 VLGNKVDKEDRQVTEEAQTCWNGDIPYLETSAKDTNVTVAEEAVRQVLAEEGLE 201

PR 14-AUG-2000; 2000US-0225759.
PR 16-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241789.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 08-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465460/50.
DR N-PSDB; AAS27054.
XX
PT Novel polypeptides useful for diagnosing, treating, preventing and/or
PT prognosing disorders related to the proteins, including cancers, immune
PT disorders and neuronal disorders -
XX
PS Claim 1; SEQ ID No 702; 880pp; English.
XX
CC The invention relates to novel isolated polypeptides (I), and
CC polynucleotides (II). (I), (II) and the antibody to (I) are useful for
CC diagnosing, preventing and treating diseases including immune system
CC disorders (e.g. congenital and acquired immunodeficiencies, autoimmune
CC disorders (e.g. rheumatoid arthritis), inflammatory conditions, organ
CC transplant rejections and graft versus host disease, infectious diseases
CC (e.g. hepatitis C), bleeding disorders, haemoglobin abnormalities and
CC other blood-related disorders (sickle cell anaemia), myeloproliferative
CC disorders, primary haematopoietic disorders, hyperproliferative
CC disorders (e.g. Gaucher's disease and cancer), neurodegenerative
CC abnormalities (Down syndrome), ischaemic injury (e.g. stroke), renal
CC disorders (e.g. glomerulonephritis), cardiovascular disorders
CC (e.g. arrhythmia), respiratory disorders, dermatological disorders
CC wound healing, epithelial cell proliferation, endocrine disorders (e.g.
CC Addison's disease), reproductive system disorders, gastrointestinal
CC disorder (inflammatory disorders), liver disorders (cirrhosis),
CC as stimulators of B-cell responsiveness to pathogens, activators of
CC T-cells, to induce higher affinity antibodies, and as a means to induce
CC tumour proliferation in pathologies e.g. acquired immune deficiency
CC syndrome (AIDS). AAU17059-AAU17683 represent novel signal transduction
CC pathway protein, amino acid sequences of the invention.
XX

Best Local Similarity 75.5%; Pred. No. 3.5e-79;
Matches: 151; Conservative 22; Mismatches 26; Indels 1; Gaps 1;

QY 1 MACKSILFKVLLGGGKSSLMNRYVTKPTOLPHIGVEFLNKDLEVDGHHFTWQI 60
Db 22 MSGKSLLLKVLGGGKSSLMNRYVTKPTOLPHIGVEFLNKDLEVDGHHFTWQI 81
QY 61 WDTAGQERFSLRTPFYRGSDCLLTFSDVDSQFNLNWKKEFTYYADVKPEPSFPFV 120
Db 82 WDTAGQERFSLRTPFYRGSDCLLTFSDVDSQFNLNWKKEFTYYADVKPEPSFPFV 141
QY 121 ILGNKIDISERQVSTEAQWCDNGDYPYFETSADQATNVAAPFEAVRRVLATEDRSD 180
Db 142 VLGNKYDKEDQVTTTEAQTWCNMGDYPYLETSAKDNTNVTVAPEAVRQVLAVEEQLE 201
QY 181 HLIQTDTVNLHRKPKDSSSC 200
Db 202 HCMLGHTIDL-TGLQAGSSC 220

RESULT 14
AAGS4480
ID AAGS4480 standard; Protein; 201 AA.
XX AAGS4480;
AC
XX
DT 18-OCT-2000 (first entry)
XX
DE
XX
KW Zea mays protein fragment SEQ ID NO: 69470.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000BP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.

```
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 50.4%; Score 535; DB 21; Length 201;
Best Local Similarity 51.5%; Pred. No. 2.6e-49;
Matches 102; Conservative 35; Mismatches 57; Indels 4; Gaps 2;

QY 7 LFKVILLGDGGVGSLSMRYVTKKEDTOLFTHTIGVEELNKDLEVDGHFVFWQIWDTAGQ 66
Db 2 LLKVIILGDSGVGKTSLMNQYNNRFSNQYKATIGADFLTKVEQIDRLFTQIWDTAGQ 61
QY 67 ERFRSLRTPFYRGSDCCLLTFSVDDSQSFQNLNNKXKEFIYYADVKPESPFPVILGNKI 126
Db 62 ERFQSLGVAFYRGADCCVLVYDVNKSKEKLNWREEFLLQASPSDPENFPFVILGNKI 121
QY 127 DI---SERQVSTEEAQWCDNGDYFYFETSAKDATNVAAPPEAVRRVLATEDRSD-HL 182
Db 122 DVDGNSRTVSEKKAKAWCASKGNIPYFETSAKEGPNVEAFECTIARNVAKNEPEEDMYL 181
QY 183 IQDTVNLHKKPKPSSSC 200
Db 182 PDTIDVGAGRQQRSSGC 199

RESULT 15
AAG54479
ID AAG54479 standard; Protein; 207 AA.
XX AAG54479;
AC AAG54479;
XX 18-OCT-2000 (first entry)
XX Zea mays protein fragment SEQ ID NO: 69469.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; Genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX Zea mays subsp. mays.
XX EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
```

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 14-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.


```
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      50.4%; Score 535; DB 21; Length 207;
Best Local Similarity 51.5%; Pred. No. 2.7e-49;
Matches 102; Conservative 35; Mismatches 57; Indels 4; Gaps 2;

QY 7 LFKVILLGSGGVKSLMNRYYTNKFDLTQFTTIGVEFLNKDLEVDGHFVTMQIWDTAGQ 66
Db 7 LFKVILLGSGGVKSLMNRYYTNKFDLTQFTTIGVEFLNKDLEVDGHFVTMQIWDTAGQ 66
QY 8 LKVIILGDSGVGKTSLSMNYNNRFSNQYKATIGADFLTKVEQIDDLFLQIWDTAGQ 67
Db 8 LKVIILGDSGVGKTSLSMNYNNRFSNQYKATIGADFLTKVEQIDDLFLQIWDTAGQ 67
QY 67 ERFRLRTPFYRGSDCLLTFPSVDDSQSFQNLNWKKEFYIYADVKEPESPPFVILGNKI 126
Db 67 ERFRLRTPFYRGSDCLLTFPSVDDSQSFQNLNWKKEFYIYADVKEPESPPFVILGNKI 126
QY 68 ERFQSLGVAFYRGADCCVLVYDVNXXSKSFEKLNWREEFLLIQASPSDPENFPFVVLGNKI 127
Db 68 ERFQSLGVAFYRGADCCVLVYDVNXXSKSFEKLNWREEFLLIQASPSDPENFPFVVLGNKI 127
QY 127 DI---SERQVSTEEAQCWRDNGDYPFETSADATNVAARFEEAVRRVLATEDRSD-HL 182
Db 127 DI---SERQVSTEEAQCWRDNGDYPFETSADATNVAARFEEAVRRVLATEDRSD-HL 182
QY 128 DVDGNSRTVSEKKAKAWCASKGNIPYFETSAGKGFNVAAFECAIARNVAVNEPEEDMYL 187
Db 128 DVDGNSRTVSEKKAKAWCASKGNIPYFETSAGKGFNVAAFECAIARNVAVNEPEEDMYL 187
QY 183 IOTDTVNLHRLKPKPSSC 200
Db 183 IOTDTVNLHRLKPKPSSC 200
QY 188 PDTIDVGGAGRQQRSSGC 205
Db 188 PDTIDVGGAGRQQRSSGC 205
```

Search completed: February 10, 2004, 04:55:52
Job time : 77 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 04:54:36 ; Search time 27 Seconds
(without alignments)
314.981 Million cell updates/sec

Title: US-09-988-974-5

Perfect score: 1062

Sequence: 1 MAGKSLFKVILLGDGVGK.....LIQDTVNLHRKRPSSCC 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/prodata/2/aaa/5A COMB pep:*
- 2: /cgn2_6/prodata/2/aaa/5B COMB pep:*
- 3: /cgn2_6/prodata/2/aaa/6A COMB pep:*
- 4: /cgn2_6/prodata/2/aaa/6B COMB pep:*
- 5: /cgn2_6/prodata/2/aaa/PCTUS COMB pep:*
- 6: /cgn2_6/prodata/2/aaa/backfiles1 pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1062	100.0	201	2	US-08-773-423-5 Sequence 5, Appli
2	1057	99.5	201	2	US-08-773-423-9 Sequence 9, Appli
3	538.5	50.7	202	2	US-08-531-525-24 Sequence 24, Appl
4	538.5	50.7	202	2	US-08-718-270A-24 Sequence 24, Appl
5	527	49.6	203	2	US-08-531-525-23 Sequence 23, Appl
6	527	49.6	203	2	US-08-718-270A-23 Sequence 23, Appl
7	374	35.2	214	2	US-08-531-525-52 Sequence 52, Appl
8	374	35.2	214	2	US-08-718-270A-52 Sequence 52, Appl
9	335.5	31.6	201	2	US-08-531-525-13 Sequence 13, Appl
10	335.5	31.6	201	2	US-08-718-270A-13 Sequence 13, Appl
11	333.5	31.4	202	2	US-08-531-525-14 Sequence 14, Appl
12	333.5	31.4	202	2	US-08-718-270A-14 Sequence 14, Appl
13	327	30.8	217	2	US-08-773-423-3 Sequence 3, Appli
14	325.5	30.6	203	2	US-08-766-551-8 Sequence 8, Appli
15	323	30.4	213	2	US-08-531-525-11 Sequence 11, Appl
16	323	30.4	213	2	US-08-718-270A-11 Sequence 11, Appl
17	322	30.3	201	2	US-08-916-901-3 Sequence 3, Appli
18	322	30.3	201	4	US-08-154-603-3 Sequence 3, Appli
19	322	30.3	207	2	US-08-824-873-4 Sequence 4, Appli
20	322	30.3	207	3	US-08-198-184-4 Sequence 4, Appli
21	321	30.2	201	2	US-08-916-901-8 Sequence 8, Appli
22	321	30.2	201	4	US-09-154-602-8 Sequence 8, Appli
23	320	30.1	213	2	US-08-773-423-8 Sequence 8, Appli
24	318	29.9	204	4	US-09-149-476-363 Sequence 363, App
25	317	29.8	215	2	US-08-531-525-10 Sequence 10, Appl
26	317	29.8	215	2	US-08-718-270A-10 Sequence 10, Appl
27	311	29.3	207	2	US-08-531-525-35 Sequence 35, Appl

Sequence 35, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 36, Appl
Sequence 3, Appl
Sequence 12, Appl
Sequence 12, Appl
Sequence 51, Appl
Sequence 67, Appl
Sequence 9, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 16, Appl
Sequence 16, Appl

ALIGNMENTS

RESULT 1

US-08-773-423-5

; Sequence 5, Application US/08773423

; Patent No. 5869291

; GENERAL INFORMATION:

; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Goli, Surya K.

; APPLICANT: Bandman, Olga

; TITLE OF INVENTION: NOVEL RAB PROTEINS

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/773,423

; FILING DATE: Herewith

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0183 US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 201 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: Consensus

; CLONE: Consensus

US-08-773-423-5

Query Match 100.0%; Score 1062; DB 2; Length 201;

Best Local Similarity 100.0%; Pred. No. 1.8e-112;

Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGKSSLFKVLILGDGCGKSSLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHFVTMQI 60
Db 1 MAGKSSLFKVLILGDGCGKSSLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHFVTMQI 60
Qy 61 WDTAGQERFSLRTPPFYRGSDCCLLTFSDVDSQSFQNLNWKKEFIYYADVKEPSFPFV 120
Db 61 WDTAGQERFSLRTPPFYRGSDCCLLTFSDVDSQSFQNLNWKKEFIYYADVKEPSFPFV 120
Qy 121 ILGNKIDISROVSTBEAQAWCRDNGDYPYFETSAKDNTNVAFAFEAVRVLATEDRSD 180
Db 121 ILGNKIDISROVSTBEAQAWCRDNGDYPYFETSAKDNTNVAFAFEAVRVLATEDRSD 180
Qy 181 HLIQDTVNLHRKPKSSSCC 201
Db 181 HLIQDTVNLHRKPKSSSCC 201

RESULT 2
US-08-773-423-9
; Sequence 9, Application US/08773423
; Patent No. 5869291
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; APPLICANT: Bandman, Olga
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,423
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION NUMBER:
; APPLICATION DATA:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0183 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 486830
US-08-773-423-9

Query Match 99.5%; Score 1057; DB 2; Length 201;
Best Local Similarity 99.5%; Pred. No. 6.8e-112;
Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGKSSLFKVLILGDGCGKSSLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHFVTMQI 60

Db 1 MAGKSSLFKVLILGDGCGKSSLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHFVTMQI 60
Qy 61 WDTAGQERFSLRTPPFYRGSDCCLLTFSDVDSQSFQNLNWKKEFIYYADVKEPSFPFV 120
Db 61 WDTAGQERFSLRTPPFYRGSDCCLLTFSDVDSQSFQNLNWKKEFIYYADVKEPSFPFV 120
Qy 121 ILGNKIDISROVSTBEAQAWCRDNGDYPYFETSAKDNTNVAFAFEAVRVLATEDRSD 180
Db 121 ILGNKIDISROVSTBEAQAWCRDNGDYPYFETSAKDNTNVAFAFEAVRVLATEDRSD 180
Qy 181 HLIQDTVNLHRKPKSSSCC 201
Db 181 HLIQDTVNLHRKPKSSSCC 201

RESULT 3
US-08-531-525-24
; Sequence 24, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
US-08-531-525-24

Query Match 50.7%; Score 538.5; DB 2; Length 202;
Best Local Similarity 51.2%; Pred. No. 4.6e-53;
Matches 104; Conservative 38; Mismatches 58; Indels 3; Gaps 3;

Qy 1 MAGKSSLFKVLILGDGCGKSSLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHFVTMQI 60
Db 1 MTKKKVLLKVLILGDGCGKTSLMNQYVNNKFNQYKATIGADFLTKELMWDVDRVVTMQI 60
Qy 61 WDTAGQERFSLRTPPFYRGSDCCLLTFSDVDSQSFQNLNWKKEFIYYADVKEPSFPFV 120

Db 61 WDTAQERFSLRGVAFYRGADCCVLCYDNNVAKTFENLDSWRDEFLIQAGPRDPNPFV 120
QY 121 ILGNKIDI-SRQVSTEEQAQWCRDNGDYPYFETSAKDATTNVAAPFEAVRVLATEDRS 179
Db 121 VLGNKIDLENQVWSQGAASQSKGNIPYFETSAKEAINVEQAFQTIARNAIKLEDGL 180
QY 180 DHLIOTDTVNLHRKPKPS-SSCC 201
Db 181 VFPIPTN-IQVPEPQPAKSGCC 202

RESULT 4

US-08-718-270A-24
; Sequence 24, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Dictyostelium discoideum
US-08-718-270A-24

Query Match 50.7%; Score 538.5; DB 2; Length 202;
Best Local Similarity 51.2%; Pred. No. 4.6e-53;
Matches 104; Conservative 30; Mismatches 58; Indels 3; Gaps 3;
QY 1 MAGKSLFKVLLGDGGVGSLSLNRYVYTKFPTQLFHTTGVEFLNKDLEVDGHFVTMQI 60
Db 1 MTKKKVLLKVLIGDSGVGKTSLSNQYVYTKFPTQLFHTTGVEFLNKDLEVDGHFVTMQI 60

QY 61 WDTAQERFSLRTPFYRGSDCCLLTFSDSQSFQNLNNKKEFIYYADVKEPESFPFV 120
Db 61 WDTAQERFQSLGVAFYRGADCCVLCYDNNVAKTFENLDSWRDEFLIQAGPRDPNPFV 120
QY 121 ILGNKIDI-SRQVSTEEQAQWCRDNGDYPYFETSAKDATTNVAAPFEAVRVLATEDRS 179
Db 121 VLGNKIDLENQVWSQGAASQSKGNIPYFETSAKEAINVEQAFQTIARNAIKLEDGL 180
QY 180 DHLIOTDTVNLHRKPKPS-SSCC 201
Db 181 VFPIPTN-IQVPEPQPAKSGCC 202

RESULT 5

US-08-531-525-23
; Sequence 23, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-531-525-23

Query Match 49.6%; Score 527; DB 2; Length 203;
Best Local Similarity 51.5%; Pred. No. 9.3e-52;
Matches 102; Conservative 30; Mismatches 64; Indels 2; Gaps 2;

QY 4 KSSLFKVLLGDGGVGSLSLNRYVYTKFPTQLFHTTGVEFLNKDLEVDGHFVTMQIWD 63
Db 5 KKVLLKVLIGDSGVGKTSLSNQYVYTKFPTQLFHTTGVEFLNKDLEVDGHFVTMQIWD 64
QY 64 AQGERFSLRTPFYRGSDCCLLTFSDSQSFQNLNNKKEFIYYADVKEPESFPFVILG 123
Db 65 AQGERFQSLGV-FYRGADCCVLCYDNNVAKTFENLDSWRDEFLIQAGPRDPNPFVILG 123

QY 124 NKIDISERQVSTEEAQACWCRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSDDL 183
Db 124 NKIDLENQVATKRAQAWCYSKNNIPYFETSAKEAINVEQAFQTTIARNALKOETEVELYN 183
QY 184 Q-TDTVNLHRKPKPSSC 200
Db 184 EPPEIKLKDKAATSAEC 201

RESULT 6

US-08-718-270A-23
; Sequence 23, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Canis familiaris
US-08-718-270A-23

Query Match 49.6%; Score 527; DB 2; Length 203;
Best Local Similarity 51.5%; Pred. No. 9.3e-52;
Matches 102; Conservative 30; Mismatches 64; Indels 2; Gaps 2;

QY 4 KSSLFKVILLGGVGKSLMRYVTKPDQTLFTIGVEFLNKDLEVDGHVFTMQIWD 63
Db 5 KVVLLKVVILGGVGKSLMRYVTKPDQTLFTIGVEFLNKDLEVDGHVFTMQIWD 64
QY 64 AQQERFSLRTFFYRGSDCCLLTFFSVDDSQSFQNLNNKKEFIYYADVKEPESFPFVLIG 123

Db 65 AQQERFSLGV-FYRGADCCVLVDFVTAENTFKTLDSWRDEFLIOASPRDPENFPFVLIG 123
QY 124 NKIDISERQVSTEEAQACWCRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSDDL 183
Db 124 NKIDLENQVATKRAQAWCYSKNNIPYFETSAKEAINVEQAFQTTIARNALKOETEVELYN 183
QY 184 Q-TDTVNLHRKPKPSSC 200
Db 184 EPPEIKLKDKAATSAEC 201

RESULT 7

US-08-531-525-52
; Sequence 52, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-531-525-52

Query Match 35.2%; Score 374; DB 2; Length 214;
Best Local Similarity 38.8%; Pred. No. 2.3e-34;
Matches 80; Conservative 44; Mismatches 66; Indels 16; Gaps 6;

QY 7 LFKVILLGGVGKSLMRYVTKPDQTLFTIGVEFLNKDLEVDGHVFTMQIWD 66
Db 11 LFKVILLGGVGKSLMRYVTKPDQTLFTIGVEFLNKDLEVDGHVFTMQIWD 70
QY 67 ERFRSLRTFFYRGSDCCLLTFFSVDDSQSFQNLNNKKEFIYYADVKEPESFPFVLGNKI 126
Db 71 ERYRAITSAYRGAVGALLVYDIKHLTYENVERMLKELRDHAD----SNIVIMLVGNKS 126
QY 127 DISE-ROVSTEEAQACWCRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSDDL- 181

Db 127 DLRLRAVPTDEARAFKNG-LSFIETGALDSTNVAAAFQILTEIYRIYVSKQMSDRE 185
Qy 182 ---LIQTDVTLNRKP---KPSSSCC 201
Db 186 NDMSPSNVPIHVPTTEKPKVQCC 211

RESULT 8
US-08-718-270A-52
; Sequence 52, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; TITLE OF INVENTION: the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 52:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 214 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-718-270A-52

Query Match 35.2%; Score 374; DB 2; Length 214;
Best Local Similarity 38.8%; Pred. No. 2.3e-34;
Matches 80; Conservative 44; Mismatches 66; Indels 16; Gaps 6;

Qy 7 LFKVILGDSGVKSLMRYVTNKFDTLFHTIGVEFLNKDLVDGHHFVTMOWDTAGQ 66
Db 11 LFKVILGDSGVKSLMRYVTNKFDTLFHTIGVEFLNKDLVDGHHFVTMOWDTAGQ 70

Qy 67 ERFRLRTPFYRGSDCCLLTFVSDDSQSQFNLNNKKFYYADYVKEPSPFPVILGNKI 126
Db 71 ERYRAITSAYRGAVGALLVYDIKHLTYENVERWIKELRDHAD----SNIVIMLVGNKS 126

Qy 127 DISE-ROVSTEEACAMCRDNGDYFYEETSAKDATNVAAFE---EAVRVLATEDRSDH- 181
Db 127 DLRLRAVPTDEARAFKNG-LSFIETGALDSTNVAAAFQILTEIYRIYVSKQMSDRE 185

Qy 182 ---LIQTDVTLNRKP---KPSSSCC 201
Db 186 NDMSPSNVPIHVPTTEKPKVQCC 211

RESULT 9
US-08-531-525-13
; Sequence 13, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; TITLE OF INVENTION: of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discoptye ommata
; US-08-531-525-13

Query Match 31.6%; Score 335.5; DB 2; Length 201;
Best Local Similarity 36.9%; Pred. No. 4.9e-30;
Matches 76; Conservative 40; Mismatches 67; Indels 23; Gaps 6;

Qy 7 LFKVILGDSGVKSLMRYVTNKFDTLFHTIGVEFLNKDLVDGHHFVTMOWDTAGQ 66
Db 8 LFKVILGDSGVKSLMRYVTNKFDTLFHTIGVEFLNKDLVDGHHFVTMOWDTAGQ 67

Qy 67 ERFRLRTPFYRGSDCCLLTFVSDDSQSQFNLNNKKFYYADYVKEPSPFPVILGNKI 126
Db 68 ERFRLRTPFYRGSDCCLLTFVSDDSQSQFNLNNKKFYYADYVKEPSPFPVILGNKI 123

Qy 127 DISE-ROV---STEEACAMCRDNGDYFYEETSAKDATNVAAFEAVRVL-----ATED 177
Db 124 DLTKKVVDTTKE-----FADSLGIFLETSAKNATNVQAFMTMAEIKRMGPGATSG 179

QY 178 RSDHLIQTDTVNLHRKPKSS--SCC 201
Db 180 GS-----EKSNNVNIQSTPVKSSGGGCC 201

RESULT 10

US-08-718-270A-13
; Sequence 13, Application US/08718270A
; Patent No. 5910478
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5910478le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptidomimetics Inhibiting
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/718,270A
; FILING DATE: 20-SEP-1996
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/531,525
; FILING DATE: 21-SEP-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,091
; FILING DATE: 21-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 78-95
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 201 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Discopyge omata
US-08-718-270A-13

Query Match 31.6%; Score 335.5; DB 2; Length 201;
Best Local Similarity 36.9%; Pred. No. 4.9e-30;
Matches 76; Conservative 40; Mismatches 67; Indels 23; Gaps 6;
QY 7 LFKVILLGGGVGKSLMRYVTKFDTLQFHTIGVEFLNKLQLEVDGHHFVWQIWDTAGQ 66
Db 8 LFKLLIIGDSGVGKSLLRFADDTYTESYISTIGVDFKIRTIELDKGTIKLQIWDTAGQ 67
QY 67 ERFRSLRTPFYRGSCDCLLTFVSDSQSPQNLNNWKEFIYVADVKEPESFPFVILGNKI 126
Db 68 ERFRITTSYRGAHGIIVYDVTDQESFNWVQWLOEIDRYAS-----ENVNKLIVGNKC 123
QY 127 DISERQVSTEEAQAQCRDNGDVPYFETSADATNVAAPFEAVRVL-----ATED 177

Db 124 DLTTKKVVDVTTKE----PADSLGIFLETSAKATNVEQAFWMTAAEIKRKGPGNTSG 179
QY 178 RSDHLIQTDTVNLHRKPKSS--SCC 201
Db 180 GS-----EKSNNVNIQSTPVKSSGGGCC 201

RESULT 11

US-08-531-525-14
; Sequence 14, Application US/08531525
; Patent No. 5840683
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; the Oncogenic Action of P21 Ras
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 202 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Lymnea stagnalis
US-08-531-525-14

Query Match 31.4%; Score 333.5; DB 2; Length 202;
Best Local Similarity 34.7%; Pred. No. 8.3e-30;
Matches 70; Conservative 46; Mismatches 69; Indels 17; Gaps 4;
QY 7 LFKVILLGGGVGKSLMRYVTKFDTLQFHTIGVEFLNKLQLEVDGHHFVWQIWDTAGQ 66
Db 11 LFKLLIIGDSGVGKSLLRFADDTYTESYISTIGVDFKIRTIELDKGTIKLQIWDTAGQ 70
QY 67 ERFRSLRTPFYRGSCDCLLTFVSDSQSPQNLNNWKEFIYVADVKEPESFPFVILGNKI 126
Db 71 ERFRITTSYRGAHGIIVYDVTDQESFNWVQWLOEIDRYAS-----ENVNKLIVGNKS 126
QY 127 DISERQVSTEEAQAQCRDNGDVPYFETSADATNVAAPFEAVRVLATEDS 179
Db 127 DLTTKKVVDVTTKE----PADSLGIFLETSAKATNVEQAFWMTAAEIKRKGPGNTSG 185
QY 180 DLIQTDTVNLHRKPKSSCC 201

RESULT 14
US-08-766-551-8
; Sequence 8, Application US/08766551
; Patent No. 5840569
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Hawkins, Phillip R.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN GTP-BINDING PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,551
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0168 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 203 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: 1457955
; US-08-766-551-8

Query Match 30.6%; Score 325.5; DB 2; Length 203;
Best Local Similarity 38.9%; Pred. No. 6.8e-29;
Matches 77; Conservative 39; Mismatches 73; Indels 9; Gaps 5;

Qy 7 LFKVILLGGGVGKSLMRYVYTNKEDTQLFHTIGVEFLNKLDEVDGHEFTVMQIWDTAGQ 66
Db 9 LFKVILGNAGVGKTCVLRFTQGLPPGQATIGVGFMIKTVGEINGEKVKLIQIWDTAGQ 68

Qy 67 ERSRLTPFYRGSDCLITFSVDDSQSNLSNWKKEFYIYADVKEPESFFVILGNKI 126
Db 69 ERFPSITQSYRSANALITDITCSFRCLPFWLREIEQVASNK----VITVLVGNKI 124

Qy 127 DISE-RQVSTEEAQAQCRNGDYPFETSAKDATNVAAAFEEAVRRVLATEDRSDLIQT 185
Db 125 DLAEERVSQRAEESQAQDMY-VLETSKESDNVKEFLDLACR-LISEARQNTLVNN 182

Qy 186 DTVNLRKPKPS--SCC 201
Db 183 VSSPLPGEKGSISYLTC 200

RESULT 15
US-08-531-525-11

; Sequence 11, Application US/08531525
; Patent No. 5840883
; GENERAL INFORMATION:
; APPLICANT: Hlavka, Joseph J.
; APPLICANT: Pincus, Matthew R.
; APPLICANT: No. 5840683le, John F.
; APPLICANT: Abajian, Henry B.
; APPLICANT: Kende, Andrew S.
; TITLE OF INVENTION: Peptides Inhibiting the Oncogenic Action
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C.
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: Colorado
; COUNTRY: US
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/531,525
; FILING DATE: 21-SEP-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33,878
; REFERENCE/DOCKET NUMBER: 37-94
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 499-8080
; TELEFAX: (303) 499-8089
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 213 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; US-08-531-525-11

Query Match 30.4%; Score 323; DB 2; Length 213;
Best Local Similarity 41.0%; Pred. No. 1.4e-28;
Matches 73; Conservative 35; Mismatches 58; Indels 12; Gaps 5;

Qy 1 MAGKSS-----LRFKVVLLGGGVGKSLMRYVYTNKEDTQLFHTIGVEFLNKLDEVDGHP 55
Db 1 MAGYADSEYDYLFLKVLIGDSGVGKSNLSRF-TQNFNLESKSTIGVEFATKTKVEGV 59

Qy 56 VTQMIDTAGQERFRLTPFYRGSDCLITFSVDDSQSNLSNWKKEFYIYADVKEPE 115
Db 60 VKAQIWDTAGQERVRAITSAYRGAVALIYDTRHATENAARWLRELGHDT----P 115

Qy 116 SFPFVILGNKIDISER-QVSTEEAQAQCRNGDYPFETSAKDATNVAAAFEEAVRRV 172
Db 116 NIVVMVLGNKCDLRLHVAVKTEAKAFERESLY-FMETSALDATNVENAFTEVLTOI 172

Search completed: February 10, 2004, 04:59:42
Job time : 29 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 04:56:01 ; Search time 75 Seconds
(without alignments)
561.145 Million cell updates/sec

Title: US-09-988-974-5

Perfect score: 1062

Sequence: 1 MAGKSLFKVILLGDBGVKK.....LIQDTNVLHLPKPPSSCC 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 801455

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	201	9	US-09-988-974-5
2	1062	100.0	209	9	US-09-988-974-5
3	1057	99.5	221	10	US-09-988-974-5
4	839	79.0	202	10	US-09-988-974-5
5	812.5	76.5	221	10	US-09-988-974-5
6	529	49.8	208	9	US-09-925-302-534
7	524.5	49.4	206	9	US-09-925-302-534
8	518.5	48.8	209	12	US-10-369-493-598
9	495	46.6	208	12	US-10-369-493-598
10	473.5	44.6	250	12	US-10-369-493-598
11	439.5	41.4	267	12	US-10-369-493-598
12	423.5	39.9	270	12	US-10-369-493-598
13	419.5	39.5	198	9	US-09-925-302-534
14	419.5	39.5	198	10	US-09-925-302-534
15	419.5	39.5	198	10	US-09-925-302-534

16	341	32.1	211	12	US-10-369-493-598
17	341	32.1	223	12	US-10-369-493-598
18	336	31.6	201	12	US-10-369-493-598
19	336	31.6	201	12	US-10-369-493-598
20	334.5	31.5	204	15	US-10-128-714-8214
21	333	31.4	207	12	US-10-369-493-598
22	332	31.3	207	9	US-09-925-302-534
23	332	31.3	246	9	US-09-925-302-534
24	331.5	31.2	207	12	US-10-369-493-598
25	329.5	31.0	222	12	US-10-369-493-598
26	327	30.8	190	10	US-09-925-302-534
27	327	30.8	213	15	US-10-369-493-598
28	327	30.8	217	9	US-09-988-974-5
29	326	30.7	241	12	US-10-369-493-598
30	324.5	30.6	205	12	US-10-369-493-598
31	324	30.5	222	10	US-09-925-302-534
32	323.5	30.5	215	12	US-10-369-493-598
33	323	30.4	217	15	US-10-369-493-598
34	323	30.4	239	9	US-09-925-302-534
35	322	30.3	201	10	US-09-925-302-534
36	322	30.3	201	12	US-10-369-493-598
37	322	30.3	221	12	US-10-369-493-598
38	322	30.3	224	15	US-10-102-806-466
39	321.5	30.3	225	15	US-10-128-714-8214
40	321	30.2	201	10	US-09-925-302-534
41	320.5	30.2	222	12	US-10-369-493-598
42	320	30.1	213	9	US-09-988-974-5
43	319.5	30.1	206	15	US-10-128-714-8214
44	319.5	30.1	206	15	US-10-128-714-8214
45	319.5	30.1	210	12	US-10-369-493-598

ALIGNMENTS

RESULT 1

; Sequence 5, Application US/09988974
; Patent No. US20020090712A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Goli, Surya K.
; Bandman, Olga

TITLE OF INVENTION: NOVEL RAB PROTEINS

NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/988,974

FILING DATE: 19-No. US20020090712A1-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/215,887

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0183 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:
 LENGTH: 201 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 IMMEDIATE SOURCE:
 LIBRARY: Consensus
 CLONE: Consensus
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 US-09-988-974-5

Query Match 100.0%; Score 1062; DB 9; Length 201;
 Best Local Similarity 100.0%; Pred. No. 1.1e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGKSSLFKVLILGDDGGVGSLLMRYVTKNFKDTQLFHTTIGVFLNKDLEVDGHFTVMOI 60
 DB 1 MAGKSSLFKVLILGDDGGVGSLLMRYVTKNFKDTQLFHTTIGVFLNKDLEVDGHFTVMOI 60
 QY 61 WDTAGQERFSLRTPFFRGSDCCLLTFSVDDSQSFQNLNWKKEFIYYADVKEPESFPFV 120
 DB 61 WDTAGQERFSLRTPFFRGSDCCLLTFSVDDSQSFQNLNWKKEFIYYADVKEPESFPFV 120
 QY 121 ILGNKIDISERQVSTEEAQAQWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRSD 180
 DB 121 ILGNKIDISERQVSTEEAQAQWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRSD 180
 QY 181 HLIQTDVNLHRKPKSSSCC 201
 DB 181 HLIQTDVNLHRKPKSSSCC 201

RESULT 2
 US-09-864-761-42996
 ; Sequence 42996, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Aeomica-x-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: PCT/US01/00670
 PRIOR FILING DATE: 2001-01-30
 PRIOR APPLICATION NUMBER: US 60/234,687
 PRIOR FILING DATE: 2000-09-21
 PRIOR APPLICATION NUMBER: US 09/608,408
 PRIOR FILING DATE: 2000-06-30
 PRIOR APPLICATION NUMBER: US 09/774,203
 PRIOR FILING DATE: 2001-01-29
 NUMBER OF SEQ ID NOS: 49117
 SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 SEQ ID NO 42996
 LENGTH: 209
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 OTHER INFORMATION: MAP TO AC003037.1
 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: BE735344.1, EVALUE 4.00e-82
 OTHER INFORMATION: SWISSPROT HIT: P51151, EVALUE 1.00e-108
 US-09-864-761-42996

Query Match 100.0%; Score 1062; DB 9; Length 209;
 Best Local Similarity 100.0%; Pred. No. 1.2e-106;
 Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAGKSSLFKVLILGDDGGVGSLLMRYVTKNFKDTQLFHTTIGVFLNKDLEVDGHFTVMOI 60
 DB 9 MAGKSSLFKVLILGDDGGVGSLLMRYVTKNFKDTQLFHTTIGVFLNKDLEVDGHFTVMOI 68
 QY 61 WDTAGQERFSLRTPFFRGSDCCLLTFSVDDSQSFQNLNWKKEFIYYADVKEPESFPFV 120
 DB 69 WDTAGQERFSLRTPFFRGSDCCLLTFSVDDSQSFQNLNWKKEFIYYADVKEPESFPFV 128
 QY 121 ILGNKIDISERQVSTEEAQAQWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRSD 180
 DB 129 ILGNKIDISERQVSTEEAQAQWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRSD 188
 QY 181 HLIQTDVNLHRKPKSSSCC 201
 DB 189 HLIQTDVNLHRKPKSSSCC 209

RESULT 3
 US-09-988-974-9
 ; Sequence 9, Application US/09889974
 ; Patent No. US20020090712A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hillman, Jennifer L.
 ; Goli, Surya K.
 ; Bandman, Olga
 ; TITLE OF INVENTION: NOVEL RAB PROTEINS
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.
 ; STREET: 3174 Porter Drive
 ; CITY: Palo Alto
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: DOS
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/988,974
 ; FILING DATE: 19-No. US20020090712A1-2001
 ; CLASSIFICATION: <Unknown>

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/215,887
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0183 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX: <Unknown>
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 201 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: GenBank
/ CLONE: 486830
/ SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-988-974-9

Query Match 99.5%; Score 1057; DB 9; Length 201;
Best Local Similarity 99.5%; Pred. No. 3.9e-106;
Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSSLFKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 60
DB 1 MAGKSSLFKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 60
QY 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120
DB 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120
QY 121 ILGNKIDISERQVSTEEAQAACRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
DB 121 ILGNKIDISERQVSTEEAQAACRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSSCC 201
DB 181 HLIQTDTVNLHRKPKPSSSCC 201

RESULT 4
US-09-764-868-1121
; Sequence 1121, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1121
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-868-1121

Query Match 79.0%; Score 839; DB 10; Length 222;
Best Local Similarity 76.1%; Pred. No. 1.8e-82;
Matches 153; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

QY 1 MAGKSSLFKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 60
DB 22 MSGKSLLLKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 81
QY 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120
DB 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120

RESULT 5
US-09-764-868-702
; Sequence 702, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ32
; CURRENT APPLICATION NUMBER: US/09/764,868
; PRIOR FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 702
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (221)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-868-702

Query Match 76.5%; Score 812.5; DB 10; Length 221;
Best Local Similarity 75.5%; Pred. No. 1.3e-79;
Matches 151; Conservative 22; Mismatches 26; Indels 1; Gaps 1;

QY 1 MAGKSSLFKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 60
DB 22 MSGKSLLLKVLGGGKSSLMNRYVTKFDLTQFHTTIGVFLNKLQLEVDGHHFTVMTQI 81
QY 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120
DB 61 WDTAGQERFRLRTPFYRGSDCCLLTFSVDDSQSFQNLNKKKEFYIYADVKEPSPFPV 120
QY 121 ILGNKIDISERQVSTEEAQAACRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
DB 121 ILGNKIDISERQVSTEEAQAACRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSSCC 200
DB 181 HLIQTDTVNLHRKPKPSSSCC 200

RESULT 6
US-09-925-302-629
; Sequence 629, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; CURRENT APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 629
; LENGTH: 208
```


QY 123 GNKIDISERO--VSTEEAQAACRDNGDYPFETSAKDNTVAAPFEAVRVLATERSD 180
Db 125 GNKIDAESKVLVSKEQELAKSLGDIPLFITSAGNAINTDTAFETARSALQOQADT 184
QY 181 HLIQ---TDTVNLHKKPSSSCC 201
Db 185 EAFEDDYNDAINIRLDGENNSCSC 208

RESULT 10

US-10-369-493-2572
; Sequence 2572, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2572
; LENGTH: 250
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(250)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-2572

Query Match 44.6%; Score 473.5; DB 12; Length 250;

Best Local Similarity 39.2%; Pred. No. 7.6e-43;
Matches 98; Conservative 39; Mismatches 64; Indels 49; Gaps 4;

QY 1 MAGKSS-LFKVILGDGGVGKSSLMNR-----YVTKPDTQ 35
Db 1 MAGKXHLKVLIGESGVGKTSIMNQVFXNMLMLSLCKXKLLKSVYVNRKESKD 60
QY 36 LPHIGVEFLNKDLEVDGHFTMQIDTAGQERFSLRTPPYRGSDCCLLTFSVDDSQSF 95
Db 61 YKATIGADFLTKVLDVDDKVVTLQMDTAGQERFQSLGVAFYRGADCCVLVDVNNKSF 120
QY 96 QNLGNWKEFTYVADVKEPSPFPVILGNKIDISER-----Q 132
Db 121 ETLDSWRDEFLIQASPNPETFPIILGNKVDVEQKRWCVXSIEFLIQSKBFLANNFR 180
QY 133 VSTEEAQAACRDNGDYPFETSAKDNTVAAPFEAVRVLATERSDHILQ-TDTVNLH 191
Db 181 FSKKALAFCAQGEIPIYFETSAKEINVQAEFETVAKALENMDSDIAADFTDPIHLD 240
QY 192 RKPSPSSSCC 201
Db 241 MESQKTSCYC 250

RESULT 11

US-10-369-493-12690
; Sequence 12690, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12690
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(267)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12690

Query Match 41.4%; Score 439.5; DB 12; Length 267;
Best Local Similarity 36.7%; Pred. No. 4e-39;
Matches 98; Conservative 36; Mismatches 58; Indels 75; Gaps 6;

QY 10 VILGDGGVGKSSLMNRVYT-----NKFDTLPHHTIGVEFLNKD 48
Db 1 VILGDGGVGKTSLMNQYVSYLLFPSPRAFSAHTDPTRVNKKFSGSYKATIGADFLTKE 60
QY 49 LEVDGHFTVMTQI-----WDTAQERFSLRTPPYRGSDCCLLTFSVD 90
Db 61 VLVDRLVTQARIHGTACSIHKLTKWSRSWDTAQERFQSLGVAFYRGADCCVLVDVN 120
QY 91 DSQSFQNLNWKKEFTYVADVKEPSPFP-----VILGNKIDI--SERQ 132
Db 121 NSKSFELDXWRDEFLIQASPRDPSFPVSRIGPWLRSANAFVXVIGNKIDMEESKRM 180
QY 133 VSTEEAQAACRDNGDYPFETSAKDNTVAAPFE-----EAVRVLAT 175
Db 181 ISKRAMTFQSKGNIPYFETSAKEINVQAEFISFOXYCTGTLLISAIVIARSALAQ 240
QY 176 EBRSDH-LIQTDTVNLHKKPSSSCC 201
Db 241 EBAEYGGDYTDPIINHDTTERDGCAC 267

RESULT 12

US-10-369-493-3332
; Sequence 3332, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 3332
; LENGTH: 270
; TYPE: PRT
; ORGANISM: Neurospora crassa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(270)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-3332

Query Match 39.9%; Score 423.5; DB 12; Length 270;
Best Local Similarity 34.8%; Pred. No. 2.2e-37;

Db 61 FRALRPAYRGAQGFLLVVDITSRDSFENVKKWLEEILRHAD--KDENVPILVGNKODL 118
QY 129 SE-----ROVSTEEAQWCRDNGDYPYPETSAKDATNVAAAPEEAVRRVLATED 177
Db 119 EDEDELETEGOKRVVSTEEGEALAKELGALPFFMETSAKTNTNVEEAFELAREILKK-- 176
QY 178 RSDHLIQDITVNLHR-KPKPSSCC 201
Db 177 -----VSEVNVNLDQFAKKKKSKCC 196

Search completed: February 10, 2004, 05:01:14
Job time : 79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 04:46:06 ; Search time 37 seconds
(without alignments)
522.430 Million cell updates/sec

Title: US-09-988-974-5

Perfect score: 1062

Sequence: 1 MAGKSSLFKVLGGVGVK.....LIQDTVNLHKKPKSSSCC 201

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 76:*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1062	100.0	201	2	G02361
2	1057	99.5	201	2	G36187
3	533	50.2	206	2	JC4107
4	529	49.8	207	2	B30413
5	529	49.8	207	2	JC5288
6	527	49.6	207	2	S62733
7	526.5	49.6	201	2	S01934
8	523.5	49.3	205	2	S36368
9	523	49.2	206	2	S33531
10	520.5	49.0	207	2	T12579
11	518.5	48.8	209	2	T26119
12	515	48.5	205	2	T03629
13	514	48.4	205	2	T40425
14	514	48.4	206	2	T03628
15	509	47.9	206	2	H96562
16	501.5	47.2	206	2	C96529
17	499.5	47.0	206	2	S39566
18	499	47.0	203	2	T00770
19	498	46.9	204	2	T03630
20	496.5	46.8	206	2	S39567
21	495	46.6	208	2	A44334
22	493	46.4	208	2	T04019
23	484.5	45.6	230	2	C84606
24	373	35.1	216	2	JN0056
25	373	35.1	216	2	E36364
26	373	35.1	216	2	J47169
27	369.5	34.8	216	2	JC4108
28	368	34.7	214	2	S10026
29	359.5	33.9	202	2	S38740

30	353	33.2	221	2	H71444	GTP-binding protei
31	352.5	33.2	203	2	S34253	GTP-binding protei
32	350	33.0	214	2	J86641	GTP-binding protei
33	349	32.9	218	2	C38625	GTP-binding protei
34	348	32.8	217	2	A86230	hypothetical prote
35	347	32.7	219	2	G84723	probable RAS type
36	346	32.6	218	2	JC2487	GTP-binding protei
37	346	32.6	218	2	A55005	GTP-binding protei
38	345	32.5	215	2	T06443	GTP-binding protei
39	344.5	32.4	217	2	S36365	GTP-binding protei
40	344	32.4	208	2	A38202	GTP-binding protei
41	343	32.3	202	2	S41430	GTP-binding protei
42	343	32.3	214	2	T14566	GTP-binding protei
43	341	32.1	211	2	T29035	hypothetical prote
44	341	32.1	223	2	S42679	GTP-binding protei
45	340.5	32.1	258	2	B86153	ARA-5 [imported] -

ALIGNMENTS

RESULT 1

G02361

Small GTP binding protein Rab9 - human

C/Species: Homo sapiens (man)

C/Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 02-Feb-2001

C/Accession: G02361

R/Joannou, Y.A.; Davies, J.P.

submitted to the EMBL Data Library, December 1995

A/Reference number: H01115

A/Accession: G02361

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-201 <IOA>

A/Cross-references: EMBL:U4103; NID:g1174146; PIDN:AACS1200.1; PID:g1174147

C/Superfamily: ras transforming protein; translation elongation factor Tu homology

C/Keywords: GTP binding; nucleotide binding; P-loop

F:8-127/Domain: translation elongation factor Tu homology <ETU>

F:14-21/Region: nucleotide-binding motif A (P-loop)

F:124-127/Region: GTP-binding NKXD motif

F:154-156/Region: GTP-binding SAK/L motif

Query Match		100.0%;	Score 1062;	DB 2;	Length 201;
Best Local Similarity		100.0%;	Pred. No. 1.6e-89;		
Matches 201;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MAGKSSLFKVLGGVGVKSSLMNRYVTNKFDLTQLPHTTIGVEFLNKDLEVDGHFVTMQI	60
DB	1	MAGKSSLFKVLGGVGVKSSLMNRYVTNKFDLTQLPHTTIGVEFLNKDLEVDGHFVTMQI	60
QY	61	WTAGQERFSLRTPFYRGSQCLLTFSVDDSQSFOINLNNKKKPIYYADYKEPESFFPV	120
DB	61	WTAGQERFSLRTPFYRGSQCLLTFSVDDSQSFOINLNNKKKPIYYADYKEPESFFPV	120
QY	121	ILGNKIDISERQVSTEEAQAQWCRDNGDYYPFETSAKDATNVAFAFEAVRVLATEQRSD	180
DB	121	ILGNKIDISERQVSTEEAQAQWCRDNGDYYPFETSAKDATNVAFAFEAVRVLATEQRSD	180
QY	181	HLIQDTVNLHKKPKSSSCC	201
DB	181	HLIQDTVNLHKKPKSSSCC	201

RESULT 2

S36187

GTP-binding protein rab9 - dog

C/Species: Canis lupus familiaris (dog)

C/Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 02-Feb-2001

C/Accession: S36187; C36364; S15605; S37290

R/Lombardi, D.; Soldati, T.; Riederer, M.A.; Goda, Y.; Zerial, M.; Pfeiffer, S.R.

EMBO J. 12, 677-682, 1993

A/Title: Rab9 functions in transport between late endosomes and the trans Golgi network

A/Reference number: S36187; MUID:93178443; PMID:8440258

A:Accession: S36187
A:Molecule type: mRNA
A:Residues: 1-201 <LON>
A:Cross-references: EMBL:X56386
R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.
A:Reference number: A36364; MUID:91061765; PMID:2123294
A:Accession: C36364
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 30-201 <CHA>
A:Cross-references: GB:X56386; NID:G2189; PIDN:CAA39797.1; PID:G2190
C:Genetics:
A:Gene: rab9
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:8-127/Domain: translation elongation factor Tu homology <ETU>
F:14-21/Region: nucleotide-binding motif A (P-loop)
F:124-127/Region: GTP-binding NKXD motif
F:154-156/Region: GTP-binding SAK/L motif

Query Match 99.5%; Score 1057; DB 2; Length 201;
Best Local Similarity 99.5%; Pred. No. 4.5e-89;
Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAGKSLFKVILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 60
Db 1 MAGKSLFKVILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 60

Qy 61 WDTAGGERPRLTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFV 120
Db 61 WDTAGGERPRLTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFV 120

Qy 121 ILGNKIDISERQVSTEEAQACWRDNGDYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
Db 121 ILGNKIDISERQVSTEEAQACWRDNGDYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180

Qy 181 HLIQTDTVNLH--RKPKSSSCC 201
Db 181 HLIQTDTVNLH--RKPKSSSCC 201

RESULT 3
JC4107
membrane vesicle transport protein ypt C5 - Chlamydomonas reinhardtii
N:Alternate names: ras-like yptC5 protein; small G protein yptC5
C:Species: Chlamydomonas reinhardtii
C:Date: 21-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 02-Feb-2001
C:Accession: JC4107
R:Dietmaier, W.; Fabry, S.; Huber, H.; Schmitt, R.
Gene 158, 41-50, 1995
A:Title: Analysis of a family of ypt genes and their products from Chlamydomonas reinhardtii
A:Reference number: JC4105; MUID:95309723; PMID:7789809
A:Accession: JC4107
A:Molecule type: mRNA
A:Residues: 1-206 <DIE>
A:Cross-references: GB:U13170; NID:G806725; PIDN:AA82728.1; PID:G806726
C:Comment: This protein plays an essential role in the regulation of intracellular membrane
C:Genetics:
A:Gene: yptC5
A:Introns: 18/1; 60/3; 109/3; 136/3; 185/1
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:37-45/Domain: effector #status predicted <EPF>
F:125-128/Region: GTP-binding NKXD motif
F:158-160/Region: GTP-binding SAK/L motif

Query Match 50.2%; Score 533; DB 2; Length 206;
Best Local Similarity 51.7%; Pred. No. 3e-41;
Matches 105; Conservative 33; Mismatches 59; Indels 5; Gaps 3;

Qy 4 KSSLFKVILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 63
Db 5 KRRLLKVIILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 64

Qy 64 AGQERFSLRTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFVILG 123
Db 65 AGQERFSLRTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFVILG 124

Qy 124 NKIDI---SERQVSTEEAQACWRDNGDYPYFETSAKDNTNVAAPFEAVRRVLATEDRSD 180
Db 125 NKIDENGSGSRQVSEKAKAWCASKSGIPYFETSAKEDINVEAFTCITRNALRNE--KEE 183

Qy 181 HLIQTDTVNLH--RKPKSSSCC 201
Db 184 ELFPDVAIDMNTTATQRRAGCC 206

RESULT 4
B30413
GTP-binding protein rab7 - dog
C:Species: Canis lupus familiaris (dog)
C:Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 02-Feb-2001
C:Accession: B30413; C39648
R:Chavrier, P.; Vingron, M.; Sander, C.; Simons, K.; Zerial, M.
Mol. Cell. Biol. 10, 6578-6585, 1990
A:Title: Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell line.
A:Reference number: A36364; MUID:91061765; PMID:2123294
A:Accession: B30413
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-207 <CHA>
A:Cross-references: GB:M35522; NID:G164057; PIDN:AAA30890.1; PID:G164058
R:Chavrier, P.; Parton, R.G.; Hauri, H.P.; Simons, K.; Zerial, M.
Cell 62, 317-329, 1990
A:Title: Localization of low molecular weight GTP binding proteins to exocytic and endo
A:Reference number: A39648; MUID:90322428; PMID:2115402
A:Accession: C39648
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-207 <CH2>
A:Cross-references: GB:M35522; PIDN:AAA30890.1; PID:G164058
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:125-128/Region: GTP-binding NKXD motif
F:155-157/Region: GTP-binding SAK/L motif

Query Match 49.8%; Score 529; DB 2; Length 207;
Best Local Similarity 56.6%; Pred. No. 7.1e-41;
Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

Qy 4 KSSLFKVILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 63
Db 5 KRVLLKVIILLGDGSGVKSLLMNRVYTNKFTQTLFHTIGVEFLNKDLEVDGHHFTVMTQI 64

Qy 64 AGQERFSLRTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFVILG 123
Db 65 AGQERFSLRTPPYRSGDCCLLTFSDSOFNLNWKKEFIYADVKEPSPFVILG 124

Qy 124 NKIDISERQVSTEEAQACWRDNGDYPYFETSAKDNTNVAAPFEAVRRVLATE 176
Db 125 NKIDLENRQVATKRAQACWCYKKNIPYFETSAKEINVEAQFTIARNALKQE 177

RESULT 5
JC5268
GTPase Rab7 protein - human
C:Species: Homo sapiens (man)
C:Date: 04-Sep-1998 #sequence_revision 04-Sep-1998 #text_change 02-Feb-2001
C:Accession: JC5268
R:Vitelli, R.; Chiariello, M.; Lattero, D.; Bruni, C.B.; Bucci, C.

Biochem. Biophys. Res. Commun. 229, 887-890, 1996
A:Title: Molecular cloning and expression analysis of the human Rab7 GTP-ase complementary
A:Reference number: JCS268; MUID:97115674; PMID:8954989
A:Accession: JCS268
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-207 <VIT>
A:Cross-references: EMBL:X93499; NID:g1089892; PIDN:CAA63763.1; PID:g1089893
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:125-128/Region: GTP-binding NKXD motif
F:155-157/Region: GTP-binding SAK/L motif

Query Match 49.8%; Score 529; DB 2; Length 207;
Best Local Similarity 56.6%; Pred. No. 7.1e-41;
Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

Qy 4 KSSLFKVILLGDGGVGSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWD 63
Db 5 KVVLLKVIILGDSGVGKTSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWD 64

Qy 64 AGQRRSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 123
Db 65 AGQRRSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 124

Qy 124 NKIDISQVSTEAQAWCRDNGDYPFETSAKDNTVAFAAEVRRVLATE 176
Db 125 NKIDLENQVATKRAQAWCYKNNIPYFETSAKEINVEQAFQTIARNALKQE 177

RESULT 6
S62733
small GTP-binding protein Rab7 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 02-Feb-2001
C:Accession: S62733
R:Vitelli, R.; Chiarillo, M.; Bruni, C.B.; Bucci, C.
Biochim. Biophys. Acta 1264, 268-270, 1995
A:Title: Cloning and expression analysis of the murine Rab7 cDNA.
A:Reference number: S62733; MUID:96138545; PMID:8547311
A:Accession: S62733
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-207 <VIT>
A:Cross-references: EMBL:X89650; NID:g1050550; PIDN:CAA61797.1; PID:g1050551
C:Genetics:
A:Gene: Rab7
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:125-128/Region: GTP-binding NKXD motif
F:155-157/Region: GTP-binding SAK/L motif

Query Match 49.6%; Score 527; DB 2; Length 207;
Best Local Similarity 56.6%; Pred. No. 1.1e-40;
Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

Qy 4 KSSLFKVILLGDGGVGSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWD 63
Db 5 KVVLLKVIILGDSGVGKTSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWD 64

Qy 64 AGQRRSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 123
Db 65 AGQRRSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 124

Qy 124 NKIDISQVSTEAQAWCRDNGDYPFETSAKDNTVAFAAEVRRVLATE 176
Db 125 NKIDLENQVATKRAQAWCYKNNIPYFETSAKEINVEQAFQTIARNALKQE 177

RESULT 7
S01934
GTP-binding protein, 23K - rat
N:Alternate names: ras-related protein, 23K
C:Species: Rattus norvegicus (Norway rat)
C:Date: 18-Oct-1989 #sequence_revision 18-Oct-1989 #text_change 02-Feb-2001
C:Accession: S01934
R:Bucci, C.; Frunzio, R.; Chiarillo, L.; Brown, A.L.; Rechler, M.M.; Bruni, C.B.
Nucleic Acids Res. 16, 9979-9993, 1988
A:Title: A new member of the ras gene superfamily identified in a rat liver cell line.
A:Reference number: S01934; MUID:89057527; PMID:3057452
A:Accession: S01934
A:Molecule type: mRNA
A:Residues: 1-201 <BUG>
A:Cross-references: EMBL:X12535
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; membrane trafficking; nucleotide binding; P-loop
F:9-122/Domain: translation elongation factor Tu homology <ETU>
F:9-16/Region: nucleotide-binding motif A (P-loop)
F:119-122/Region: GTP-binding NKXD motif
F:149-151/Region: GTP-binding SAK/L motif

Query Match 49.6%; Score 526.5; DB 2; Length 201;
Best Local Similarity 50.5%; Pred. No. 1.2e-40;
Matches 100; Conservative 32; Mismatches 63; Indels 3; Gaps 1;

Qy 7 LFKVILLGDGGVGSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWDTAGQ 66
Db 2 LKVVILLGDGGVGSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWDTAGQ 61

Qy 67 ERFSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 126
Db 62 ERFSLRTPFYRGSDCCLLTFSDSOFNLNWKKEFTYADVKEPSPFPVILG 121

Qy 127 DISERQVSTEAQAWCRDNGDYPFETSAKDNTVAFAAEVRRVLATE 183
Db 122 DLENQVATKRAQAWCYKNNIPYFETSAKEINVEQAFQTIARNALKQE 181

Qy 184 QTDVNLHRKPKPSSSCC 201
Db 182 EPIKDKNERAKAESC 199

RESULT 8
S36368
GTP-binding protein yptV5 - Volvox carteri
C:Species: Volvox carteri
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 02-Feb-2001
C:Accession: S36368
R:Fabry, S.; Jacobsen, A.; Huber, H.; Palme, K.; Schmitt, R.
Curr. Genet. 24, 229-240, 1993
A:Title: Structure, expression, and phylogenetic relationships of a family of ypt genes
A:Reference number: S36368; MUID:94037148; PMID:8221932
A:Accession: S36368
A:Molecule type: DNA
A:Residues: 1-205 <FAB>
A:Cross-references: GB:L08131; NID:g409167; PIDN:AAA34254.1; PID:g409168
C:Genetics:
A:Gene: yptV5
A:Introns: 18/2; 60/3; 109/3; 135/3; 184/2
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:61-67/Region: GTP binding #status predicted
F:122-129/Region: GTP binding #status predicted
F:150-156/Region: GTP binding #status predicted

Query Match 49.3%; Score 523.5; DB 2; Length 205;
Best Local Similarity 50.5%; Pred. No. 2.2e-40;
Matches 102; Conservative 34; Mismatches 61; Indels 5; Gaps 3;

Qy 4 KSSLFKVILLGDGGVGSLSMRYVTKFTQFHTIGVFLNKDLEVDGHFVTMQIWD 63

Db 5 KRRVLKIIILGDSGVGKTSLNQVYQKFTKEYKATIGADPLTKIEIVDDKKVTMQLNDT 64
QY 64 AGQERFSLRTPFFRGSDCLLTSTVSDSQFNLNWKKEFIYADVKEPESFPFVILG 123
Db 65 AGQERFQSLGAFYRGADCCMLVFDVNNAKSFDDLNRDEFIQAGFSPDFNPFVILG 124
QY 124 NKIDI---SERQVSTEEAQAACRNDGYPYFETSADKATNVAAAFEEAVRRVLATEDRSDH 181
Db 125 NKIDVGNVRQVTEKAKAWCASKGSPYFETSADKEDINVERAFTCTRNALRNE-KEEE 183
QY 182 LIQTDVTNLRHK--PKPSSSCC 201
Db 184 LFVPDAVDMNTSATQRKRGCC 205

RESULT 9
S33531
C:Species: Pisum sativum rab - garden pea
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 02-Feb-2001
C:Accession: S33531; S25543
R:Drew, J.B.; Bown, D.; Gatehouse, J.A.
Plant Mol. Biol. 21, 1195-1199, 1993
A:Title: Sequence of a novel plant ras-related cDNA from Pisum sativum.
A:Reference number: S33531; MUID:93257636; PMID:8490139
A:Accession: S33531
A:Molecule type: mRNA
A:Residues: 1-206 <DRE>
A:Cross-references: EMBL:X65650; NID:q20755; PIDN:CAA46600.1; PID:q20756
A:Note: the authors translated the codon AAA for residue 48 as Thr, and CCC for residue 49
C:Genetics:
A:Gene: rab
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; nucleotide binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>
F:15-22/Region: nucleotide-binding motif A (P-loop)
F:125-128/Region: GTP-binding NKXD motif
F:158-160/Region: GTP-binding SAK/L motif

Query Match 49.2%; Score 523; DB 2; Length 206;
Best Local Similarity 48.3%; Pred. No. 2.5e-40;
Matches 97; Conservative 42; Mismatches 58; Indels 4; Gaps 2;
QY 4 KSSLFKVILLGDGGVGKSSLMNRYVYTNKFDTLQFHTIGVEFLNKLQLEVDGHFVTMQLNDT 63
Db 5 RRTLLKVIILGDSGVGKTSLNQVYQKFTKEYKATIGADPLTKIEIVDDKKVTMQLNDT 64
QY 64 AGQERFSLRTPFFRGSDCLLTSTVSDSQFNLNWKKEFIYADVKEPESFPFVILG 123
Db 65 AGQERFQSLGAFYRGADCCMLVFDVNNAKSFDDLNRDEFIQAGFSPDFNPFVILG 124
QY 124 NKIDI---SERQVSTEEAQAACRNDGYPYFETSADKATNVAAAFEEAVRRVLATEDRSD 180
Db 125 NKIDVGNVRQVTEKAKAWCASKGSPYFETSADKEDINVERAFTCTRNALRNE-KEE 183
QY 181 LIQTDVTNLRHKPKPSSSCC 201
Db 184 ELYLPDIDVGNSSQPRSTGC 204

RESULT 10
T12579
GTP-binding protein Rab7a - common ice plant
C:Species: Mesembryanthemum crystallinum (common ice plant)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jan-2000
C:Accession: T12579
R:Michalowski C.B.; Bohnert, H.J.
Submitted to the EMBL Data Library, January 1997
A:Description: Sequence of a GTP-binding protein from Mesembryanthemum crystallinum.
A:Reference number: Z17539
A:Accession: T12579
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA
A:Residues: 1-207 <MICS>
A:Cross-references: EMBL:U87142; NID:gi842068; PID:gi842069
C:Superfamily: ras transforming protein; translation elongation factor Tu homology
C:Keywords: GTP binding; P-loop
F:9-128/Domain: translation elongation factor Tu homology <ETU>

Query Match 49.0%; Score 520.5; DB 2; Length 207;
Best Local Similarity 48.1%; Pred. No. 4.2e-40;
Matches 98; Conservative 41; Mismatches 58; Indels 5; Gaps 3;
QY 4 KSSLFKVILLGDGGVGKSSLMNRYVYTNKFDTLQFHTIGVEFLNKLQLEVDGHFVTMQLNDT 63
Db 5 RRTLLKVIILGDSGVGKTSLNQVYQKFTKEYKATIGADPLTKIEIVDDKKVTMQLNDT 64
QY 64 AGQERFSLRTPFFRGSDCLLTSTVSDSQFNLNWKKEFIYADVKEPESFPFVILG 123
Db 65 AGQERFQSLGAFYRGADCCMLVFDVNNAKSFDDLNRDEFIQAGFSPDFNPFVILG 124
QY 124 NKIDI---SERQVSTEEAQAACRNDGYPYFETSADKATNVAAAFEEAVRRVLATEDRSD 180
Db 125 NKIDVGNVRQVTEKAKAWCASKGSPYFETSADKEDINVERAFTCTRNALRNE-KEE 184
QY 181 LIQTDVTNLRHKPKP-SSSCC 201
Db 185 TYL-PTDIDMAGSTRPQSSSAC 205

RESULT 11
T26119
Hypothetical protein W03C9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T26119
R:Gregory, J.; Ainscough, R.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z20155
A:Accession: T26119
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-209 <WIL>
A:Cross-references: EMBL:Z66516; PIDN:CAA91357.1; GSPDB:GNO0020; CESP:W03C9.3
A:Experimental source: clone W03C9
C:Genetics:
A:Gene: CESP:W03C9.3
A:Map position: 2
A:Introns: 39/3; 61/3; 112/2; 181/3
C:Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 48.8%; Score 518.5; DB 2; Length 209;
Best Local Similarity 48.8%; Pred. No. 6.5e-40;
Matches 101; Conservative 43; Mismatches 56; Indels 7; Gaps 4;
QY 1 MAG--KSSLFKVILLGDGGVGKSSLMNRYVYTNKFDTLQFHTIGVEFLNKLQLEVDGHFVTM 58
Db 1 MSGTRKALLKVIILGDSGVGKTSLNQVYQKFTKEYKATIGADPLTKIEIVDDKKVTMQLNDT 60
QY 59 QIWDTAGQERFSLRTPFFRGSDCLLTSTVSDSQFNLNWKKEFIYADVKEPESFP 118
Db 61 QIWDTAGQERFQSLGAFYRGADCCMLVFDVNNAKSFDDLNRDEFIQAGFSPDFNPFVILG 120
QY 119 FVILGNKIDI--SERQVSTEEAQAACRNDGYPYFETSADKATNVAAAFEEAVRRVLATE- 176
Db 121 FVILGNKVLDLSSQRAVSKRAQSWCQTKGNIPYFETSAKEALNVEAAFLAIRDALARES 180
QY 177 -DRSDHLIQTDTVNL--HRKPKPSSSC 200
Db 181 QETNDFPEFPDQIRLNPQNQQNSGC 207

RESULT 12
T03629
GTP-binding protein Rab7b - common tobacco

C;Species: Nicotiana tabacum (common tobacco)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Feb-2001
 C;Accession: T03629
 R;Haizel, T.; Merkle, T.; Turck, F.; Nagy, F.
 Plant Physiol. 108, 59-67, 1995
 A;Title: Characterization of membrane-bound small GTP-binding proteins from Nicotiana tabacum
 A;Reference number: Z14896; PMID:95303981; PMID:7784525
 A;Accession: T03629
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-205 <HA1>
 A;Cross-references: EMBL:L29275; NID:G623589; PIDN:AAA74119.1; PID:G623590
 A;Experimental source: strain SRI
 C;Genetics:
 A;Gene: Rab7b
 C;Superfamily: ras transforming protein; translation elongation factor Tu homology
 C;Keywords: GTP binding; nucleotide binding; P-loop
 F;8-127/Domain: translation elongation factor Tu homology <ETU>
 F;14-21/Region: nucleotide-binding motif A (P-loop)
 F;124-127/Region: nucleotide-binding motif A (P-loop)
 F;157-159/Region: GTP-binding SAK/L motif

Query Match 48.5%; Score 515; DB 2; Length 205;
 Best Local Similarity 46.6%; Pred. No. 1.3e-39;
 Matches 95; Conservative 44; Mismatches 61; Indels 4; Gaps 2;

QY 1 MAGKSLFKVILLGDGGVKGKSLMRYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 60
 DB 1 MFSANLVKLVILGDSGVGKTSIMNQYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 60

QY 61 WDTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 120
 DB 61 WDTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 120

QY 121 ILGNKIDI---SERQVSTEEAQWCRDNGDYFYFETSADKATNVAAAFEEAVRRVLATED 177
 DB 121 ILGNKVDIDGNSRVSEKARAWCASKGNIPYFETSADKATNVAAAFEEAVRRVLATED 177

QY 178 RSDHLIQDTVNLHRKPKPSSSCC 201
 DB 181 -EEIYLPDTLIDVGTSGQFRTGC 203

Query Match 48.4%; Score 514; DB 2; Length 206;
 Best Local Similarity 47.8%; Pred. No. 1.7e-39;
 Matches 97; Conservative 40; Mismatches 62; Indels 4; Gaps 2;

QY 2 AGKSLFKVILLGDGGVKGKSLMRYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 61
 DB 3 ARRRMLLVKLVILGDSGVGKTSIMNQYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 62

QY 62 DTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 121
 DB 63 DTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 122

QY 122 LGNKIDI---SERQVSTEEAQWCRDNGDYFYFETSADKATNVAAAFEEAVRRVLATEDR 178
 DB 123 LGNKIDVDGNSRVSEKARAWCASKGNIPYFETSADKATNVAAAFEEAVRRVLATEDR 181

QY 179 SDHLIQDTVNLHRKPKPSSSCC 201
 DB 182 EDEIVLPDTIDVAGSGQSRSTGC 204

RESULT 15
 H96562
 Hypothetical protein F19K6.10 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C;Accession: H96562
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

Db 61 LWDTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPF 120
 QY 120 VILGNKIDISE--RQVSTEEAQWCRDNGDYFYFETSADKATNVAAAFEEAVRRVLATED 177
 Db 121 ILGNKVDIDGNSRVSEKARAWCASKGNIPYFETSADKATNVAAAFEEAVRRVLATED 180
 QY 178 RSDHLIQ-DTDTVNLHRKPKPSSSCC 201
 Db 181 SDDIAADFTDPIHLDMESQRTSCYC 205

RESULT 14
 T03628
 GTP-binding protein Rab7a - common tobacco
 C;Species: Nicotiana tabacum (common tobacco)
 C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 02-Feb-2001
 C;Accession: T03628
 R;Haizel, T.; Merkle, T.; Turck, F.; Nagy, F.
 Plant Physiol. 108, 59-67, 1995
 A;Title: Characterization of membrane-bound small GTP-binding proteins from Nicotiana tabacum
 A;Reference number: Z14896; PMID:95303981; PMID:7784525
 A;Accession: T03628
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-206 <HA1>
 A;Cross-references: EMBL:L29274; NID:G623587; PIDN:AAA74118.1; PID:G623588
 A;Experimental source: strain SRI
 C;Genetics:
 A;Gene: Rab7a
 C;Superfamily: ras transforming protein; translation elongation factor Tu homology
 C;Keywords: GTP binding; nucleotide binding; P-loop
 F;9-128/Domain: translation elongation factor Tu homology <ETU>
 F;15-22/Region: nucleotide-binding motif A (P-loop)
 F;125-128/Region: GTP-binding NKXD motif
 F;158-160/Region: GTP-binding SAK/L motif

Query Match 48.4%; Score 514; DB 2; Length 206;
 Best Local Similarity 47.8%; Pred. No. 1.7e-39;
 Matches 97; Conservative 40; Mismatches 62; Indels 4; Gaps 2;

QY 2 AGKSLFKVILLGDGGVKGKSLMRYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 61
 Db 3 ARRRMLLVKLVILGDSGVGKTSIMNQYVYVTKFDTQLFHTIGVFLNKDLEVDGHFVTMQI 62

QY 62 DTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 121
 Db 63 DTAGQERFSLRTPFYRGSDCLLTFSVDSQSFNLNWKKEFYIYADVKEPSPFV 122

QY 122 LGNKIDI---SERQVSTEEAQWCRDNGDYFYFETSADKATNVAAAFEEAVRRVLATEDR 178
 Db 123 LGNKIDVDGNSRVSEKARAWCASKGNIPYFETSADKATNVAAAFEEAVRRVLATEDR 181

QY 179 SDHLIQDTVNLHRKPKPSSSCC 201
 Db 182 EDEIVLPDTIDVAGSGQSRSTGC 204

RESULT 15
 H96562
 Hypothetical protein F19K6.10 [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C;Accession: H96562
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: H96562
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <STO>
A;Cross-references: GB:AB005173; NID:g10645452; PIDN:AAG21568.1; GSPDB:GN00141
C;Genetics:
A;Gene: F19K6.10
A;Map position: 1
C;Superfamily: ras transforming protein; translation elongation factor Tu homology

Query Match 47.9%; Score 509; DB 2; Length 206;
Best Local Similarity 48.5%; Pred. No. 4.7e-39;
Matches 96; Conservative 37; Mismatches 61; Indels 4; Gaps 2;

Qy 7 LFKVILLGDGGVGVKSSLMNRYVTNKEDTQLFHTIGVEFLNKDLEVDGHFVTMQIWDTAGQ 66
Db 8 LLKVILLGDGSGVGTSLMNGFVNRKFSNQYKATIGADFLTKVEQIDDRIFTLQIWDTAGQ 67

Qy 67 ERFRSLRTPFYRGSDCCLLTFSVDDSQSNLGNWKKFIYYADVKEPESFPFVILGNKI 126
Db 68 ERFQSLGVAFYRGADCCVLVDVNVKSFONLNNWREELIQASPDPEFNPFPVILGNKT 127

Qy 127 DI---SROVSTEEAQAACRDNGDYPYFETSADATNVAAFEAVRRVLATEDRSDHLI 183
Db 128 DVDGKSRVSEKKAACAKGKNIPYFETSAKEGFNVDAFECITKNAPKNEPEEPYL 187

Qy 184 QDVTNHLRHKPKPSSSCC 201
Db 188 -PDTIDVAGGQQQRSTGC 204

Search completed: February 10, 2004, 04:59:00
Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 03:56:11 ; Search time 24 Seconds
(without alignments)
393.849 Million cell updates/sec

Title: US-09-988-974-5

Perfect score: 1062

Sequence: 1 MAGKSLFKVLLGGGVGK.....LIQTDTVNLHRKPKSSCC 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1062	100.0	201	1 RB9A HUMAN	P51151 homo sapien
2	1057	99.5	201	1 RB9A CANFA	P24408 canis faml
3	1033	97.3	201	1 RB9A RAT	Q99775 rattus norv
4	839	79.0	201	1 RB9B HUMAN	Q99775 rattus norv
5	535.5	50.4	203	1 RB7 DICDI	P36411 dictyosteli
6	533	50.2	206	1 YPT5 CHLRE	Q39573 chlamydomon
7	531	50.0	206	1 RB7 PENC	Q40787 pennisetum
8	529.5	49.9	207	1 RB7 RAT	P09527 rattus norv
9	529	49.8	207	1 RB7 CANFA	P18067 canis faml
10	529	49.8	207	1 RB7 HUMAN	P51149 homo sapien
11	527	49.6	207	1 RB7 MOUSE	P51150 mus musculu
12	524	49.3	207	1 RB7 GOSHI	Q9X988 gossypium h
13	523.5	49.3	205	1 YPT5 VOLCA	P36864 volvox cart
14	523	49.2	206	1 RB7 PEA	P31022 pisum sativ
15	520.5	49.0	207	1 RB7 MSCR	P93267 mesembryant
16	519	48.9	207	1 RB7 RABIT	Q97572 oryctolagus
17	514	48.4	205	1 YPT7 SCHPO	Q94655 schizosacch
18	509.5	48.0	205	1 RB7 NEUCR	Q9C318 neurospora
19	505	47.6	207	1 RB7 PRUAR	O24461 prunus arne
20	499	47.0	203	1 RB7 ARATH	O04157 arabidopsis
21	496.5	46.8	206	1 RB7 VIGAC	Q41640 vigna acon
22	495	46.6	208	1 YPT7 YRST	P32939 saccharomyc
23	493.5	46.5	206	1 YPT7 YRST	Q43463 glycine max
24	373	35.1	216	1 R11A HUMAN	P24410 homo sapien
25	369.5	34.8	216	1 YPT6 CHLRE	Q39572 chlamydomon
26	368	34.7	214	1 YPT3 SCHPO	P17310 schizosacch
27	359.5	33.9	202	1 R1C1 ORVSA	P40392 oryza sativ
28	350	33.0	214	1 ARAB ARATH	P28187 arabidopsis
29	349	32.9	218	1 R11B DISOM	P22129 discopyge o
30	348	32.8	217	1 RB1C ARATH	O04486 arabidopsis
31	346	32.6	218	1 R11B HUMAN	Q15907 homo sapien
32	346	32.6	218	1 R11B MOUSE	P46638 mus musculu
33	344.5	32.4	217	1 YPT2 VOLCA	P36861 volvox cart

34	344	32.4	208	1 YPT1 MAIZE	P16976 zea mays (m
35	343	32.3	214	1 RAB2 BETVU	Q39434 beta vulgar
36	342.5	32.3	218	1 R11E LOTJA	Q40195 lotus japon
37	341	32.1	223	1 YP31 YEAST	P38555 saccharomyc
38	340.5	32.1	258	1 ARAS ARATH	P28188 arabidopsis
39	337.5	31.8	203	1 YPT1 VOLCA	P31584 volvox cart
40	337.5	31.8	203	1 YPT2 MAIZE	O05737 zea mays (m
41	337	31.7	207	1 R98B RAT	P07550 rattus norv
42	336.5	31.7	218	1 YPT3 NICPL	Q01111 nicotiana p
43	336	31.6	202	1 RAB1 DISOM	P22125 discopyge o
44	334.5	31.5	217	1 R1C2 ORVSA	P40393 oryza sativ
45	333.5	31.4	216	1 ARAB ARATH	P28186 arabidopsis

ALIGNMENTS

RESULT 1					
ID	RB9A HUMAN	STANDARD;	PRT;	201 AA.	
AC	P51151,				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	15-SEP-2003	(Rel. 42, Last annotation update)			
DE	Ras-related protein Rab-9A (Rab-9).				
GN	RAB9A OR RAB9.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=97271569; PubMed=9126495;				
RA	Davies J.P., Cotter P.D., Ioannou Y.A.;				
RT	"Cloning and mapping of human Rab7 and Rab9 cDNA sequences and				
RT	identification of a Rab9 pseudogene."				
RL	Genomics 41:131-134 (1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Muscle;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Stausberg R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,				
RA	Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villaion D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield V.S.N., Krzyzinski M.I., Skalska U., Marra M.A.;				
RA	Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length				
RT	human and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).				
CC	-I- FUNCTION: INVOLVED IN THE TRANSPORT OF PROTEINS BETWEEN				
CC	THE ENDOSOMES AND THE TRANS GOLGI NETWORK.				
CC	-I- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				

```

DR EMBL; U44103; AAC51200.1; -.
DR EMBL; BC017265; AAH17265.1; -.
DR PIR; G02361; G02361.
DR HSP; P01112; IPLJ.
DR Genew; HGNC:9792; RAB9A.
DR MIM; 300284; -.
DR GO; GO:0003928; F:RAB small monomeric GTPase activity; TAS.
DR GO; GO:0006899; P:non-selective vesicle transport; TAS.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfrng.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack.
FT NP_BIND 14 21 GTP (BY SIMILARITY).
FT NP_BIND 62 66 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 36 44 EFECTOR REGION (BY SIMILARITY).
FT DOMAIN 200 200 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 201 AA; 22837 MW; 55B502C21E97DB72 CRC64;

Query Match 100.0%; Score 1062; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 3.1e-87;
Matches 201; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGDGGVGKSSLMNRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFVTMQI 60
DB 1 MAGKSLFKVILLGDGGVGKSSLMNRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFVTMQI 60
QY 61 WDTAGQERFSLRTPFYRGSDCCLLTFSDVSDSOFNLNWKKEFYIYADVKEPSFPFV 120
DB 61 WDTAGQERFSLRTPFYRGSDCCLLTFSDVSDSOFNLNWKKEFYIYADVKEPSFPFV 120
QY 121 ILGNKIDISEROVSTEAQAQWCRDNGDYPYFETSAKDATNVAAPFAEAVRRVLATEDRSD 180
DB 121 ILGNKIDISEROVSTEAQAQWCRDNGDYPYFETSAKDATNVAAPFAEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSCC 201
DB 181 HLIQTDTVNLHRKPKPSSCC 201

RESULT 2
RB9A_CANFA STANDARD; PRT; 201 AA.
AC P24408;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-9A (Rab-9).
GN RAB9A OR RAB9.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cocker spaniel;
RX MEDLINE=93178443; PubMed=8440258;
RA Lombardi D., Soldati T., Riederer M.A., Goda Y., Zerial M.,
RA Pfeiffer S.R.;
RT "Rab9 functions in transport between late endosomes and the trans
RL Golgi network";
RL ENBO J. 12:677-682(1993).
RN [2]
RP SEQUENCE OF 30-201 FROM N.A.
RC STRAIN=Cocker spaniel;
RX MEDLINE=91061765; PubMed=2123294;
RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
RA "Molecular cloning of YP1/SEC4-related cDNAs from an epithelial cell

```

```

RT line.".
RL MOL. Cell. Biol. 10:6578-6585(1990).
CC -!- FUNCTION: INVOLVED IN THE TRANSPORT OF PROTEINS BETWEEN
CC THE ENDOSOMES AND THE TRANS GOLGI NETWORK.
CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56386; CA39797.1; -.
DR PIR; S36187; S36187.
DR HSP; P01112; IPLJ.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsmfrng.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack.
FT NP_BIND 14 21 GTP (BY SIMILARITY).
FT NP_BIND 62 66 GTP (BY SIMILARITY).
FT NP_BIND 124 127 GTP (BY SIMILARITY).
FT DOMAIN 36 44 EFECTOR REGION (BY SIMILARITY).
FT DOMAIN 200 200 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 201 AA; 22810 MW; 65B502C204BEDB72 CRC64;

Query Match 99.5%; Score 1057; DB 1; Length 201;
Best Local Similarity 99.5%; Pred. No. 8.6e-87;
Matches 200; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAGKSLFKVILLGDGGVGKSSLMNRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFVTMQI 60
DB 1 MAGKSLFKVILLGDGGVGKSSLMNRYVTKNFKDTQLFHTIGVEFLNKDLEVDGHFVTMQI 60
QY 61 WDTAGQERFSLRTPFYRGSDCCLLTFSDVSDSOFNLNWKKEFYIYADVKEPSFPFV 120
DB 61 WDTAGQERFSLRTPFYRGSDCCLLTFSDVSDSOFNLNWKKEFYIYADVKEPSFPFV 120
QY 121 ILGNKIDISEROVSTEAQAQWCRDNGDYPYFETSAKDATNVAAPFAEAVRRVLATEDRSD 180
DB 121 ILGNKIDISEROVSTEAQAQWCRDNGDYPYFETSAKDATNVAAPFAEAVRRVLATEDRSD 180
QY 181 HLIQTDTVNLHRKPKPSSCC 201
DB 181 HLIQTDTVNLHRKPKPSSCC 201

RESULT 3
RB9A_RAT STANDARD; PRT; 201 AA.
AC Q99P75;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ras-related protein Rab-9A (Rab-9).
GN RAB9A OR RAB9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Bone;
RX MEDLINE=22047202; PubMed=12051767;
RA Zhao H., Ettala O., Vaananen H.K.;
RA "Intracellular membrane trafficking pathways in bone-resorbing

```


osteoclasts revealed by cloning and subcellular localization studies
 of small GTP-binding rab proteins.";
 Biochem. Biophys. Res. Commun. 293:1060-1065(2002).
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF PROTEINS BETWEEN
 CC THE ENDOSOMES AND THE TRANS GOLGI NETWORK (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF325692; AAC49586.1; -;
 DR HSP; P10114; 1KAO.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmfng.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRfams; TIGR00231; small GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack.
 FT NP_BIND 14 21 GTP (BY SIMILARITY).
 FT NP_BIND 62 66 GTP (BY SIMILARITY).
 FT NP_BIND 124 127 GTP (BY SIMILARITY).
 FT DOMAIN 36 44 EFFECTOR REGION (BY SIMILARITY).
 FT DOMAIN 200 200 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 201 AA; 22865 MW; 56E3ASAD656E1 CRC64;

 Query Match 97.3%; Score 1033; DB 1; Length 201;
 Best Local Similarity 95.5%; Pred. No. 1.2e-94;
 Matches 192; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MAGKSLFKVLLGDGSGVSSLMNRYVTNKTQTLPHFTIGVEFLNKDLEVDGHFVTMQI 60
 DB 1 MAGKSLFKVLLGDGSGVSSLMNRYVTNKTQTLPHFTIGVEFLNKDLEVDGHFVTMQI 60
 QY 61 WDTAGQERFSLRTPFYRGSDCCLLTFSVDDSDSFQNLNWKKEFIYADVKEPESFPFV 120
 DB 61 WDTAGQERFSLRTPFYRGSDCCLLTFSVDDSDSFQNLNWKKEFIYADVKEPESFPFV 120
 QY 121 ILGNKIDISERQVSTEAQAWCRDNGDYPYFETSAKDNTVAAPFAVRVLATEDRSD 180
 DB 121 ILGNKIDISERQVSTEAQAWCRDNGDYPYFETSAKDNTVAAPFAVRVLATEDRSD 180
 QY 181 HLIQTDTVNLHRRKPKPSSCC 201
 DB 181 HLIQTDTVNLHRRKPKPSSCC 201
 RESULT 4
 RB9B_HUMAN STANDARD; PRT; 201 AA.
 AC Q9NP50;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-9B (Rab-9L) (RAB9-like protein).
 GN RAB9B OR RAB9L.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=20496223; PubMed=11043518;
 RA Seki N., Azuma T., Yoshikawa T., Masuho Y., Muramatsu M., Saito T.;
 RT "cDNA cloning of a new member of the Ras superfamily, RAB9-like, on
 the human chromosome Xq22.1-q22.3 region.";

J. Hum. Genet. 45:318-322(2000).
 [2] SEQUENCE FROM N.A.
 RA Bird C.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN THE TRANSPORT OF PROTEINS BETWEEN
 CC THE ENDOSOMES AND THE TRANS GOLGI NETWORK (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AB036693; BAA89542.1; -;
 DR EMBL; AL139228; CAB76967.1; -;
 DR HSP; P10114; 1KAO.
 DR Genew; HGNC:14090; RAB9B.
 DR MIM; 300285; -;
 DR GO; GO:0005525; P-GTP binding activity; NAS.
 DR GO; GO:0006895; P-Golgi to endosome transport; NAS.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmfng.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMNG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRfams; TIGR00231; small GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport; Golgi stack.
 FT NP_BIND 14 21 GTP (BY SIMILARITY).
 FT NP_BIND 62 66 GTP (BY SIMILARITY).
 FT NP_BIND 124 127 GTP (BY SIMILARITY).
 FT DOMAIN 36 44 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 200 200 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 201 201 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 201 AA; 22719 MW; B35EDA8B8358C49C CRC64;

 Query Match 79.0%; Score 839; DB 1; Length 201;
 Best Local Similarity 76.1%; Pred. No. 1.9e-67;
 Matches 153; Conservative 22; Mismatches 26; Indels 0; Gaps 0;
 QY 1 MAGKSLFKVLLGDGSGVSSLMNRYVTNKTQTLPHFTIGVEFLNKDLEVDGHFVTMQI 60
 DB 1 MSGKSLLLKVVLLGDGSGVSSLMNRYVTNKTQTLPHFTIGVEFLNKDLEVDGRFVTLQI 60
 QY 61 WDTAGQERFSLRTPFYRGSDCCLLTFSVDDSDSFQNLNWKKEFIYADVKEPESFPFV 120
 DB 61 WDTAGQERFSLRTPFYRGSDCCLLTFSVDDSDSFQNLNWKKEFIYADVKEPESFPFV 120
 QY 121 ILGNKIDISERQVSTEAQAWCRDNGDYPYFETSAKDNTVAAPFAVRVLATEDRSD 180
 DB 121 VLGNKVDKEDQVTTTEAQTWCWENGDPYLETSAKDDTNTVTVAFAEVRQVLAVEQLE 180
 QY 181 HLIQTDTVNLHRRKPKPSSCC 201
 DB 181 HMLGHTIDLNSGSKAGSSCC 201
 RESULT 5
 RAB7_DICDI STANDARD; PRT; 203 AA.
 ID "RAB7_DICDI
 AC P36411;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab7.
 GN RAB7.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.

```

OX NCBI_TaxID=44689;
RN SEQUENCE FROM N.A.
RP STRAIN=AK3;
RC Bush J.M. IV, Nolta K., Rodriguez-Paris J., Temesvari L.,
RA Ruscetti T., Steck T., Cardelli J.A.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U02928; AAA80152.1; -.
DR HSP; P01112; IPLJ.
DR DictyDb; D000070; -.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LIPID 202 202 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 203 203 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 203 AA; 22695 MW; 8984547A7725B521 CRC64;

Query Match
Best Local Similarity 50.4%; Score 535.5; DB 1; Length 203;
Matches 103; Conservative 38; Mismatches 56; Indels 3; Gaps 3;

OY 4 KSSLFVKVILGDGGVGSLLMRYVYVTKFTQLFHTIGVEFLNKLDEVDGHVFTMQIWD 63
Db 5 KVLVKVILGDGGVGSLLMRYVYVTKFTQLFHTIGVEFLNKLDEVDGHVFTMQIWD 64
OY 64 AGQERFSLRTPYRGSDCLLTFSVDDSQSQFNLNWKKEFYIYADVKEPSPFPVILG 123
Db 65 AGQERFSLRTPYRGSDCLLTFSVDDSQSQFNLNWKKEFYIYADVKEPSPFPVILG 124
OY 124 NKIDI-SERQVSTEEAQWCRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 182
Db 125 NKIDLENQVRSQKRAWSQCKGNIPYFETSAKEINVEQAFQIARNAIKLEDGLVFP 184
OY 183 IQDTNLHRLKPKPS--SSCC 201
Db 185 IPTN-IQVPEPQPAKSGCC 203

RESULT 6
YPT5_CHLRE
ID YPT5_CHLRE STANDARD; PRT; 206 AA.
AC Q39573;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein YPT5.
GN YPT5
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonas.
OX NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cw15;

OX NCBI_TaxID=44689;
RN SEQUENCE FROM N.A.
RP STRAIN=AK3;
RC Bush J.M. IV, Nolta K., Rodriguez-Paris J., Temesvari L.,
RA Ruscetti T., Steck T., Cardelli J.A.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U13170; AAA82728.1; -.
DR PIR; JC4107; JC4107.
DR HSP; P10114; IKAQ.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small_gtp; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 206 206 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 23142 MW; 2F02E469BBDAA0A2 CRC64;

Query Match
Best Local Similarity 50.2%; Score 533; DB 1; Length 206;
Matches 105; Conservative 33; Mismatches 59; Indels 6; Gaps 3;

OY 4 KSSLFVKVILGDGGVGSLLMRYVYVTKFTQLFHTIGVEFLNKLDEVDGHVFTMQIWD 63
Db 5 KRLKLVILGDGGVGSLLMRYVYVTKFTQLFHTIGVEFLNKLDEVDGHVFTMQIWD 64
OY 64 AGQERFSLRTPYRGSDCLLTFSVDDSQSQFNLNWKKEFYIYADVKEPSPFPVILG 123
Db 65 AGQERFSLRTPYRGSDCLLTFSVDDSQSQFNLNWKKEFYIYADVKEPSPFPVILG 124
OY 124 NKIDI-SERQVSTEEAQWCRDNGDYPYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
Db 125 NKIDENGSSRQVSEKAKAWCASKGSIYPFETSAKEDINVEAAFTCTRNALRNE-KEE 183
OY 181 HLIQDTNLH--RKPKPSSCC 201
Db 184 ELFPDAVDNNTATQKRGCC 206

RESULT 7
RAB7_PENCL
ID RAB7_PENCL STANDARD; PRT; 206 AA.
AC Q40787;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ras-related protein Rab7 (Possible apospory-associated protein).
GN Rab7
OS Pennisetum glauca (Buffelgrass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Paniceae; Pennisetum.
OX NCBI_TaxID=35520;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-Higgins; TISSUE=flower;
RA Gustine D.L., Hulce D.A., Moyer B.G.;
RT "A novel cDNA encoding a rab7-related small GTP-binding protein in
RL Penicillium ciliare (buffelgrass)";
RN (In) Plant Gene Register PGR95-132.
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U40219; AAA85273.1; --
DR HSSP; P01112; 1PLJ
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfam; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Frenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT LPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
FT LPID 206 206 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; 22996 MW; 4B3B82CC9C55899A CRC64;
Query Match 50.0%; Score 531; DB 1; Length 206;
Best Local Similarity 49.3%; Pred. No. 4e-40;
Matches 100; Conservative 40; Mismatches 57; Indels 6; Gaps 3;
QY 4 KSSLFVKVLLGGGVGKSSLMNRYVYTKFTQLPHTTIGVEFLNKDLEVDGHFTVMTQWDT 63
DB 5 RTLLKVLIIIGDSGVGKTSLMNQVYVKKFNQYKATIGADFLTKVEQFEDRLFTLQWDT 64
QY 64 AGQERFSLRTPFYRGSDCLLTFSDSDSQFONLSNWKKEFIYADVKEPESFPFVLIG 123
DB 65 AGQERFQSLGVAFYRGADCCVLVFDVTAPNTFKTLDSEWDEFLIQAQSPDPENFPFVLIG 124
QY 124 NKIDISERQVSTEEAQWCRDNGDYPYFETSADATNVAAPFAEAVRVRLATE--DRSD 180
DB 125 NKVDVGGNRRVSEKKAKAWCASKGNIPYFETSKEGNTNVEDAFQCIKVNALKNEP--EE 183
QY 181 ELHIQTDVNL--HRKPKPSSSC 201
DB 184 ELYVFDVDDVVGNGRAQRSSGCC 206
RESULT 8
RAB7 RAT STANDARD; PRT; 207 AA.
AC P09527;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ras-related protein Rab-7 (RAS-related protein p23) (RAS-related
DE protein BRL-RAS).
GN RAB7.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Buffalo; TISSUE=Liver;
RX MEDLINE=89057527; PubMed=3057452;

RA Bucci C., Franzio R., Chiariotti L., Brown A.L., Rechler M.M.,
RA Bruni C.B.;
RT "A new member of the ras gene superfamily identified in a rat liver
RN cell line";
RN Nucleic Acids Res. 16:9979-9994(1988).
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: FOUND ON LATE ENDOSOMES (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X12535; CAA31053.1; --
DR EMBL; AF286535; AAG00543.1; --
DR HSSP; P01112; 1PLJ
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfam; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Frenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
FT LPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
FT LPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23504 MW; A2AF33B02F672971 CRC64;
Query Match 49.9%; Score 529.5; DB 1; Length 207;
Best Local Similarity 50.2%; Pred. No. 5.4e-40;
Matches 101; Conservative 32; Mismatches 65; Indels 3; Gaps 1;
QY 4 KSSLFVKVLLGGGVGKSSLMNRYVYTKFTQLPHTTIGVEFLNKDLEVDGHFTVMTQWDT 63
DB 5 KVLVLKVLIIIGDSGVGKTSLMNQVYVKKFNQYKATIGADFLTKVEVVDRLVTMQWDT 64
QY 64 AGQERFSLRTPFYRGSDCLLTFSDSDSQFONLSNWKKEFIYADVKEPESFPFVLIG 123
DB 65 AGQERFQSLGVAFYRGADCCVLVFDVTAPNTFKTLDSEWDEFLIQAQSPDPENFPFVLIG 124
QY 124 NKIDISERQVSTEEAQWCRDNGDYPYFETSADATNVAAPFAEAVRVRLATE--DRSD 180
DB 125 NKIDLENQVATKKAQAWCYSKNNIPYFETSKEGNTNVEDAFQCIKVNALKQETVELYN 184
QY 181 ELHIQTDVNLHRKPKPSSSC 201
DB 185 EFPEPKLDKNERAKASAES 205
RESULT 9
RAB7 CANFA STANDARD; PRT; 207 AA.
ID RAB7 CANFA
AC P18067;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-7.
 GN RAB7.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90322428; PubMed=2115402;
 RA Chavrier P., Parton R.G., Hauri H.P., Simons K., Zerial M.;
 RT "Localization of low molecular weight GTP binding proteins to
 RT exocytic and endocytic compartments.";
 RL Cell 62:317-329(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=Cocker spaniel;
 RA Chavrier P., Vingron M., Sander C., Simons K., Zerial M.;
 RT "Molecular cloning of YPT1/SEC4-related cDNAs from an epithelial cell
 RT line.";
 RL Mol. Cell. Biol. 10:6578-6585(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9126495;
 RA Davies J.P., Cotter P.D., Ioannou Y.A.;
 RT "Cloning and mapping of human Rab7 and Rab9 cDNA sequences and
 RT identification of a Rab9 pseudogene.";
 RL Genomics 41:131-134(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 RA Kim J.Y., Park Y.B.;
 RT Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: FOUND ON LATE ENDOSOMES.
 CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M35522; AAA30890.1; -;
 DR PIR; B30413; B30413.
 DR HSP; P01112; 1PLJ.
 DR InterPro; IPR003579; GTPase Rab.
 DR InterPro; IPR001806; Ras trnsfrmg.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMs; TIGR00231; small GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 125 128 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 207 AA; 23520 MW; E3AF33B16A67296D CRC64;
 Query Match 49.8%; Score 529; DB 1; Length 207;
 Best Local Similarity 56.6%; Pred. No. 6e-40;
 Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;
 QY 4 KSLFLKVLILGDSGVKSLMNYVTKPDTQLFHTIGVEFLNKDLVDGHVFWQIWDT 63
 Db 5 KVVLLKVLILGDSGVKSLMNYVTKPDTQLFHTIGVEFLNKDLVDGHVFWQIWDT 64
 QY 64 AGQERFSLRTPPYRGSDCLITFSYDSDQSFQNLNNKKKEFIYADVKEPSPFPFVILG 123
 Db 65 AGQERFQSLGAVPYRGADCCVLVFDVTAPTNFKLDSNRDEFLIQASPRDPENFFVVLG 124
 QY 124 NKIDISGRVSTEEAQAACRDNGDYFETSADATNVAAAEAVRRVLATE 176
 Db 125 NKIDLENQVATKRAQAWCYSKNNIPYFETSAKEAINVEQAFQTARNALKOE 177
 RESULT 10
 RAB7_HUMAN

ID RAB7 HUMAN STANDARD; PRT; 207 AA.
 AC P51149; Q9UPB0;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-7.
 GN RAB7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Placenta;
 RA Vitelli R., Chiariello M., Lattero D., Bruni C.B., Bucci C.;
 RT "Molecular cloning and expression analysis of the human Rab7 GTP-ase
 RT complementary deoxyribonucleic acid.";
 RL Biochem. Biophys. Res. Commun. 229:887-890(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97271569; PubMed=9126495;
 RA Davies J.P., Cotter P.D., Ioannou Y.A.;
 RT "Cloning and mapping of human Rab7 and Rab9 cDNA sequences and
 RT identification of a Rab9 pseudogene.";
 RL Genomics 41:131-134(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 RA Kim J.Y., Park Y.B.;
 RT Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: FOUND ON LATE ENDOSOMES (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X93499; CAA63763.1; -;
 DR EMBL; U44104; AAA86640.1; -;
 DR EMBL; AF050175; AAD02565.1; -;
 DR PIR; JCS268; JCS268.
 DR HSP; P01112; 1PLJ.
 DR Genew; HGNC:9788; RAB7.
 DR MIM; 602238; -;
 DR GO; GO:0005770; C:late endosome; TAS.
 DR GO; GO:0003928; F:RAB small monomeric GTPase activity; TAS.
 DR GO; GO:0006897; P:endocytosis; TAS.
 DR InterPro; IPR003579; GTPase Rab.
 DR InterPro; IPR001806; Ras trnsfrmg.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTRNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMs; TIGR00231; small GTP; 1.
 KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
 FT NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 125 128 GTP (BY SIMILARITY).
 FT DOMAIN 37 45 EFFECTOR REGION (BY SIMILARITY).
 FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 FT CONFLICT 47 47 T -> I (IN REF. 3).
 FT CONFLICT 108 108 I -> V (IN REF. 2).
 FT CONFLICT 127 127 I -> V (IN REF. 2).
 FT CONFLICT 180 180 V -> E (IN REF. 3).
 SQ SEQUENCE 207 AA; 23490 MW; A2AF33B16A672971 CRC64;

Query Match 49.8%; Score 529; DB 1; Length 207;
 Best Local Similarity 56.6%; Pred. No. 6e-40;
 Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

QY 4 KSSLFVKVLLGGGVGKSLMRYVYTNKFDLTQFHTIGVEFLNKDLEVDGHFVTMQIWD 63
 DB 5 KVLVLLKVIILGDSGVGKTSLSNQYVNNKFSNQYKATIGADFLTKVWVDDRLVTMQIWD 64

QY 64 AGQERFSRLTPFYRGSDCCLLTFSDVDSQSFONLSNWKKEFIYYADVKEPSPFVILG 123
 DB 65 AGQERFSRLTPFYRGSDCCLLTFSDVDSQSFONLSNWKKEFIYYADVKEPSPFVILG 124

QY 124 NKIDISERQVSTEEAQACWCRDNGDYPYFETSADATNVAAAFEEAVRRVLATE 176
 DB 125 NKIDLENQVATKRAQAQWCYKNNIPYFETSAKEAINVEQAFQTIARNALKQE 177

RESULT 11
 RAB7_MOUSE
 ID RAB7_MOUSE STANDARD; PRT; 207 AA.
 AC PS1150;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab-7.
 GN RAB7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96138545; PubMed=8547311;
 RA Vitelli R., Chiariello M., Bruni C.B., Bucci C.;
 RT "Cloning and expression analysis of the murine Rab7 cDNA."
 RL Biochim. Biophys. Acta 1264:268-270(1995).
 CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: FOUND ON LATE ENDOSOMES (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; X89650; CAA61797.1; -
 PIR; S62733; S62733.
 HSP; P01112; IPLJ.
 MGD; MGI:105068; Rab7.
 GO; GO:0005794; C:Golgi apparatus; IDA.
 GO; GO:0005770; C:late endosome; IDA.
 InterPro; IPR003579; GTPase Rab.
 InterPro; IPR001806; Ras trnsfrmg.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00071; ras; 1.
 PRINTS; PR00449; RASTNSFRMG.
 SMART; SM00175; RAB; 1.
 TIGRfams; TIGR00231; small GTP; 1.
 GTP-binding; Lipoprotein; Prenylation; Protein transport.
 NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 125 128 GTP (BY SIMILARITY).
 FT NP_BIND 37 45 EFECTOR REGION (BY SIMILARITY).
 FT DOMAIN 205 205 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 207 AA; A2AB82016A672870 CRC64;

Query Match 49.6%; Score 527; DB 1; Length 207;

Best Local Similarity 56.6%; Pred. No. 9e-40;
 Matches 98; Conservative 24; Mismatches 51; Indels 0; Gaps 0;

QY 4 KSSLFVKVLLGGGVGKSLMRYVYTNKFDLTQFHTIGVEFLNKDLEVDGHFVTMQIWD 63
 DB 5 KVLVLLKVIILGDSGVGKTSLSNQYVNNKFSNQYKATIGADFLTKVWVDDRLVTMQIWD 64

QY 64 AGQERFSRLTPFYRGSDCCLLTFSDVDSQSFONLSNWKKEFIYYADVKEPSPFVILG 123
 DB 65 AGQERFSRLTPFYRGSDCCLLTFSDVDSQSFONLSNWKKEFIYYADVKEPSPFVILG 124

QY 124 NKIDISERQVSTEEAQACWCRDNGDYPYFETSADATNVAAAFEEAVRRVLATE 176
 DB 125 NKIDLENQVATKRAQAQWCYKNNIPYFETSAKEAINVEQAFQTIARNALKQE 177

RESULT 12
 RAB7_GOSHI
 ID RAB7_GOSHI STANDARD; PRT; 207 AA.
 AC Q9XER8;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ras-related protein Rab7.
 GN RAB7.
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Malvales; Malvaceae; Malvoidae; Gossypium.
 CX NCBI_TaxID=3635;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Deltapine 62; TISSUE=Etisolated cotyledon;
 RA Jenkins W.B., Turley R.B., Steele M.;
 RT "A cDNA clone encoding a Rab7 protein from upland cotton (Gossypium
 hirsutum L.)."
 RL (In) Plant Gene Register PGR99-064.
 CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
 CC TRAFFIC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).

 EMBL; AF116243; AAD22451.1; -
 HSP; P36017; LEKO.
 InterPro; IPR003579; GTPase Rab.
 InterPro; IPR001806; Ras trnsfrmg.
 InterPro; IPR005225; Small_GTP.
 Pfam; PF00071; ras; 1.
 PRINTS; PR00449; RASTNSFRMG.
 SMART; SM00175; RAB; 1.
 TIGRfams; TIGR00231; small GTP; 1.
 GTP-binding; Lipoprotein; Prenylation; Protein transport.
 NP_BIND 15 22 GTP (BY SIMILARITY).
 FT NP_BIND 63 67 GTP (BY SIMILARITY).
 FT NP_BIND 125 128 GTP (BY SIMILARITY).
 FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
 FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
 SQ SEQUENCE 207 AA; 23171 MW; 2453D5D3B0CC617 CRC64;

Query Match 49.3%; Score 524; DB 1; Length 207;
 Best Local Similarity 51.0%; Pred. No. 1.7e-39;
 Matches 101; Conservative 34; Mismatches 59; Indels 4; Gaps 2;

QY 7 LFKVLLGGGVGKSLMRYVYTNKFDLTQFHTIGVEFLNKDLEVDGHFVTMQIWD 66
 DB 8 LLKVIILGDSGVGKTSLSNQYVNNKFSNQYKATIGADFLTKVWVDDRLVTMQIWD 67

QY 67 ERFRSLRTPYRGSDCCLLTFSVDDSSQFNLNWKKEFYIYADVKEPSPFPFVILGNKI 126
D 68 ERFOSGLVAFYRGADCCVLDVNNWKSFDNANWREFFLIQASPDPEPFVILGNKV 127
QY 127 DI---SERQVSTEAQACWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRS-D-HL 182
D 128 DVDGNSRVVSEKKAACWASKGNIPYFETSAKEGPNVDAFEACIARNALKNPEBEIYL 187
QY 183 IQDITVNLHKKPKPSSSC 200
D 188 PETIDVASGGRPORSTGC 205

RESULT 13
YPT5 VOLCA STANDARD; PRT; 205 AA.
AC P36864; (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE GTP-binding protein ypt5.
GN YPT5.
OS Volvox carteri.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Volvocaceae; Volvox.
OX NCBI_TaxID=3067;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=f. Nagariensis / HK10;
RX MEDLINE=94037148; PubMed=8221932;
RA Fabry S., Jacobsen A., Huber H., Palme K., Schmitt R.;
RT "Structure, expression, and phylogenetic relationships of a family of
RT YPT genes encoding small G-proteins in the green alga Volvox
RT carteri";
RL Curr. Genet. 24:229-240(1993).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L08131; AAA34254.1; -;
DR PIR; S36368; S36368.
DR HSSP; P01112; 1JAH.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport;
KW Multigene family.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT DOMAIN 37 45 EFFECTOR REGION (PROBABLE).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 205 AA; A0DD8B8764D7E163 CRC64;

Query Match 49.3%; Score 523.5; DB 1; Length 205;
Best Local Similarity 50.5%; Pred. No. 1.8e-39;
Matches 102; Conservative 34; Mismatches 61; Indels 5; Gaps 3;

QY 4 KSLFKVILLGDSGVGKSLMNRVYTNKFDTLQFHTIGVFLNKDLEVDGHVFMQIWD 63

Db 5 KRVLKIIILLGDSGVGKSLMNRVYTNKFDTLQFHTIGVFLNKDLEVDGHVFMQIWD 64
QY 64 AGQRFSLRTPYRGSDCCLLTFSVDDSSQFNLNWKKEFYIYADVKEPSPFPFVILG 123
D 65 AGQRFOSGLVAFYRGADCCVLDVNNWKSFDNANWREFFLIQASPDPEPFVILG 124
QY 124 NKIDI--SERQVSTEAQACWCRDNGDYPYFETSAKDNTVAAPFEAVRVLATEDRS-DH 181
D 125 NKIDVGNRVQVTEKKAACWASKGSIYPYFETSAKEDINVERAFTCITRNALRNE-KEEE 183
QY 182 LIQDITVNLHKK--PKPSSSC 201
D 184 LFVDAVDMNTSATQQRKGCC 205

RESULT 14
RAB7_PEA STANDARD; PRT; 206 AA.
ID RAB7_PEA
AC P31022;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ras-related protein Rab7.
OS Pisum sativum (garden pea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.
OX NCBI_TaxID=3888;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Purple-podded; TISSUE=Pod;
RX MEDLINE=93257636; PubMed=8490139;
RA Drew J.E., Bown D., Gatehouse J.A.;
RT "Sequence of a novel plant ras-related cDNA from Pisum sativum.";
RL Plant Mol. Biol. 21:1195-1199(1993).
CC -1- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X65650; CAA46600.1; -;
DR PIR; S33531; S33531.
DR HSSP; P36017; 1BK0.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
KW NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT LIPID 204 204 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 206 206 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 206 AA; A2126D26DBEBA49 CRC64;

Query Match 49.2%; Score 523; DB 1; Length 206;
Best Local Similarity 48.3%; Pred. No. 2e-39;
Matches 97; Conservative 42; Mismatches 58; Indels 4; Gaps 2;

QY 4 KSLFKVILLGDSGVGKSLMNRVYTNKFDTLQFHTIGVFLNKDLEVDGHVFMQIWD 63
D 5 RRTLLKVIILLGDSGVGKSLMNRVYTNKFDTLQFHTIGVFLNKDLEVDGHVFMQIWD 64

```
QY 64 AGQERFSLRTPFYRGSDCLLTFVSDDSQFNLNWKKEFIYADVKEPESFPFVILG 123
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 AGQERFQSLGVAFYRGADCCVVDVNSVKFNDLNWRREFLIQANPSDFENFPFVIG 124
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 124 NKIDI---SEROVSTEEAQCWCRONGDYPFYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 NKIDVGGSGRVVSEKKAKAWCMKGNIPYFETSAKOGTNVVEEAFQCIAXNAIQNEPSEE 184
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 HLIQTDVTNLRHKPKP-SSSSCC 201
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 TYL-PDTIDMAGSTRPQSSSAC 205
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 HLIQTDVTNLRHKPKPSSSSCC 201
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 184 ELYLPDITDVNSQPRSTGC 204
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

RESULT 15

```
RAB7_MESCR
ID RAB7_MESCR STANDARD; PRT; 207 AA.
AC P93267;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ras-related protein Rab7A.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE FROM N.A.
RA Michalowaki C.B., Bohnert H.J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROTEIN TRANSPORT. PROBABLY INVOLVED IN VESICULAR
CC TRAFFIC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U87142; RAB47557.1; -
DR FIR; T12579; T12579.
DR HSSP; P36017; IEK0.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras_trnsmg.
DR InterPro; IPR005225; Small_Gtp.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGRfams; TIGR00231; small GTP; 1.
KW GTP-binding; Lipoprotein; Prenylation; Protein transport.
FT NP_BIND 15 22 GTP (BY SIMILARITY).
FT NP_BIND 63 67 GTP (BY SIMILARITY).
FT NP_BIND 125 128 GTP (BY SIMILARITY).
FT LIPID 205 205 GERANYL-GERANYL (BY SIMILARITY).
FT LIPID 207 207 GERANYL-GERANYL (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23188 MW; 3B900B895A40DC7 CRC64;
```

```
Query Match 49.0%; Score 520.5; DB 1; Length 207;
Best Local Similarity 48.3%; Pred. No. 3.4e-39;
Matches 98; Conservative 41; Mismatches 58; Indels 5; Gaps 3;
QY 4 KSSLFVILLGDSGVGKSLMREYVTKEDTQLFHTIGVEFLNKLDEVDGHVFTMQIWDT 63
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 5 RRKLLKTIILGDSGVGKSLMNGFVNKFSNQYKATIGADFLTKELQFEDRLFTIQIWDT 64
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 64 AGQERFSLRTPFYRGSDCLLTFVSDDSQFNLNWKKEFIYADVKEPESFPFVILG 123
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 65 AGQERFQSLGVAFYRGADCCVVDVNSVKFNDLNWRREFLIQANPSDFENFPFVIG 124
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
QY 124 NKIDI---SEROVSTEEAQCWCRONGDYPFYFETSAKDATNVAAAFEEAVRRVLATEDRSD 180
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 NKIDVGGSGRVVSEKKAKAWCMKGNIPYFETSAKOGTNVVEEAFQCIAXNAIQNEPSEE 184
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 181 HLIQTDVTNLRHKPKP-SSSSCC 201
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 TYL-PDTIDMAGSTRPQSSSAC 205
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
Search completed: February 10, 2004, 04:56:30
Job time : 26 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2004, 04:44:31 ; Search time 85 Seconds
(without alignments)
610.219 Million cell updates/sec

Title: US-09-988-974-5

Perfect score: 1062

Sequence: 1 MAGKSLFKVILGGGVGK.....LIQTDVTLNRKPKPSSCC 201

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_23:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1034	97.4	201	11 Q9ROM6	Q9ROM6 mus musculus
2	839	79.0	201	11 Q8BHH2	Q8BHH2 mus musculus
3	531.5	50.0	205	3 Q8TGD9	Q8TGD9 emericella
4	518.5	48.8	209	5 Q23146	Q23146 caenorhabdi
5	518	48.8	206	5 Q9NFG0	Q9NFG0 plasmodium
6	515	48.5	205	10 Q40527	Q40527 nicotiana t
7	514	48.4	206	10 Q40526	Q40526 nicotiana t
8	513.5	48.4	217	10 Q8WCB7	Q8WCB7 oryza sativ
9	513	48.3	206	10 Q40213	Q40213 lotus japon
10	513	48.3	206	5 Q8LS94	Q8LS94 arabidopsis
11	512	48.2	207	5 Q76742	Q76742 drosophila
12	510.5	48.1	256	5 Q9VIM6	Q9VIM6 drosophila
13	509.5	48.0	205	3 Q9CZL8	Q9CZL8 neurospora
14	509	47.9	206	10 Q9C820	Q9C820 arabidopsis
15	508.5	47.9	206	5 Q9NHW8	Q9NHW8 entamoeba h
16	508	47.8	206	10 Q9LW76	Q9LW76 arabidopsis

17	508	47.8	209	5 Q9U5G8	Q9U5G8 tetrahymena
18	504	47.5	207	10 Q40214	Q40214 lotus japon
19	502.5	47.3	206	5 Q9BLE8	Q9BLE8 entamoeba h
20	501.5	47.2	206	10 Q9X198	Q9X198 arabidopsis
21	500	47.1	206	10 Q948X8	Q948X8 arabidopsis
22	499	47.0	206	10 Q9LGH5	Q9LGH5 arabidopsis
23	498	46.9	204	10 Q40528	Q40528 nicotiana t
24	494.5	46.6	205	10 Q40211	Q40211 lotus japon
25	493	46.4	208	10 Q9S288	Q9S288 arabidopsis
26	481	45.3	212	10 Q9SJ11	Q9SJ11 arabidopsis
27	478	45.0	169	5 Q15608	Q15608 entamoeba h
28	464.5	43.7	205	10 Q40212	Q40212 lotus japon
29	437	41.1	208	3 Q9HDY0	Q9HDY0 schizosacch
30	433.5	40.8	221	5 Q9XZK0	Q9XZK0 trypanosoma
31	428	40.3	223	5 Q8T6H9	Q8T6H9 leishmania
32	427	40.2	198	10 Q8LQJ2	Q8LQJ2 oryza sativ
33	424	39.9	223	5 Q9N2P5	Q9N2P5 leishmania
34	413	38.9	179	5 Q813B6	Q813B6 plasmodium
35	405	38.1	204	10 Q948K6	Q948K6 arabidopsis
36	398.5	37.5	183	10 Q9FK99	Q9FK99 arabidopsis
37	386.5	36.4	199	11 Q8VEA8	Q8VEA8 mus musculus
38	378.5	35.6	199	4 Q86AH8	Q86AH8 homo sapien
39	356.5	33.6	203	10 Q24112	Q24112 nicotiana p
40	353	33.2	202	10 Q9SEH3	Q9SEH3 arabidopsis
41	353	33.2	221	10 Q23594	Q23594 arabidopsis
42	352.5	33.2	203	10 Q40569	Q40569 nicotiana t
43	352.5	33.2	214	3 Q8WZR7	Q8WZR7 neurospora
44	352	33.1	203	10 Q8RU63	Q8RU63 oryza sativ
45	351	33.1	202	10 Q9FPJ4	Q9FPJ4 arabidopsis

ALIGNMENTS

RESULT 1

Q9ROM6
ID Q9ROM6 PRELIMINARY; PRT; 201 AA.
AC Q9ROM6;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Small GTP binding protein (2410064E05RUK protein) (Ssimilar to RAB9,
DE member RAS oncogene family) (RAS-related protein RAB-9 homolog).
CS RAB9 OR STD 99 GR 2410064E05RUK.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RP Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Saito T.;
RT "Mouse small GTP binding protein.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryonic stem cells, and Head;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gisi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staali F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Grotstein S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

Db 65 AGQERFQSLGVAFYRGADCCVYVYNNKSFALDSWRDEFLIQASPRDPSPFPVVG 124
Qy 124 NKID--SERQVSTEEAQCNRNGDYPYFETSAKADNTNVAAPFAEAVRRVLATEDRSDH 181
Db 125 NKIDMEESKMISSKRAMTFCQSGNTPYFETSAKEAVNVEQAFVIARSALAQEEAEY 184
Qy 182 -LIQDTVNLHRLKPKPSSSC 201
Db 185 GGDVTDPIINHDTIRDCAC 205

RESULT 4

Q23146 PRELIMINARY; PRT; 209 AA.
AC Q23146;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE W03C9.3 protein.
GN W03C9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gregory J., Ainscough R.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851915;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RT Science 282:2012-2018 (1998).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z66516; CAA91357.1; --
DR HSSP; P01112; 1PLJ.
DR WormPep; W03C9.3; CE03777.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 209 AA; 23429 MW; 1797B3E7B887F255 CRC64;

Query Match 48.8%; Score 518.5; DB 5; Length 209;
Best Local Similarity 48.8%; Pred. No. 4.4e-41;
Matches 101; Conservative 43; Mismatches 56; Indels 7; Gaps 4;
Qy 1 MAG--KSLFKVILLGDGGVKGKSLNRYVTKFDTLQFLHTIGVFLNKLVDGDFV 58
Db 1 MSSTRKALLKVIILGDSGVGKTSLMNQYVNNKFTNQYKATIGADFLTDVNDRTVL 60
Qy 59 QIWDTAGQERFSLRTPFYRGSDCLLTFSVDDSSQFQNLNWKKEFYIYADVKEPSPFP 118
Db 61 QIWDTAGQERFSLRTPFYRGSDCLLTFSVDDSSQFQNLNWKKEFYIYADVKEPSPFP 120
Qy 119 FVILGNKIDI--SRQVSTEEAQCNRNGDYPYFETSAKADNTNVAAPFAEAVRRVLATE- 176
Db 121 FVILGNKIDLESQRAVSQKRAQSCQTKGNIYYEVSKEALNVAEAFIARDALAES 180
Qy 177 -DRSDHLIQTDTNL--HRKPKPSSSC 200
Db 181 QETNDPFPDPQIRLNPQNOQNSGC 207

RESULT 5

Q9NFGO PRELIMINARY; PRT; 206 AA.
AC Q9NFGO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative Rab7 GTPase.
GN RAB7.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=21036610; PubMed=11163452;
RA Spielmann T., Beck H.P.;
RT "Analysis of stage-specific transcription in Plasmodium falciparum
RT reveals a set of genes exclusively transcribed in ring stage
RT parasites."
RL Mol. Biochem. Parasitol. 111:453-458 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RA Quevillon E., Spielmann T., Langsley G.;
RT "The rab gene family of Plasmodium falciparum";
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AJ290938; CAB92946.2; --
DR HSSP; P10114; 1XAO.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras trnsfrmg.
DR InterPro; IPR002078; Sig54 interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTNSFRMG.
DR SMART; SM00175; RAB; 1.
DR TIGRFAMs; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 206 AA; 23788 MW; E46BD3A7460084D5 CRC64;

Query Match 48.8%; Score 518; DB 5; Length 206;
Best Local Similarity 47.5%; Pred. No. 4.8e-41;
Matches 96; Conservative 42; Mismatches 60; Indels 4; Gaps 2;
Qy 4 KSLFKVILLGDGGVKGKSLNRYVTKFDTLQFLHTIGVFLNKLVDGDFV 63
Db 5 KRTILKVIILGDSGVGKTSLMNQYVNNKFTNQYKATIGADFLTKETIVDNEQITMQIWD 64
Qy 64 AQERFSLRTPFYRGSDCLLTFSVDDSSQFQNLNWKKEFYIYADVKEPSPFPVILG 123
Db 65 AQERFQSLGVAFYRGADCCVLDLTNYKTSLESWKDEFLIQASPKDPENFPFVIG 124
Qy 124 NKID--ISERQVSTEEAQCNRNGDYPYFETSAKADNTNVAAPFAEAVRRVLATEDRSDHL 182
Db 125 NKIDETNKRKVQSLKVLQWCKSNNTIPYFETSAKAINVDQAFDEIARKAKKHQEEQI 184
Qy 183 IQDTVNLHRLKPKP---SSSC 201
Db 185 YLPETPALNNQSEQKMYKSCC 206

RESULT 6

Q40527 PRELIMINARY; PRT; 205 AA.
AC Q40527;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Nt-rab7b protein.
GN NT-RAB7B.
OS Nicotiana tabacum (Common tobacco).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SRI;
 RX MEDLINE=95303981; PubMed=7784525;
 RA Haizel T., Merkle T., Turck F., Nagy F.;
 RT "Characterization of membrane-bound small GTP-binding proteins from
 RT Nicotiana tabacum."
 RL Plant Physiol. 108:59-67(1995).
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; L29275; AAA74119.1; -.
 DR HSSP; P36017; 1EK0.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 205 AA; 22744 MW; 596A6F1BA68A7D1B CRC64;

Query Match 48.5%; Score 515; DB 10; Length 205;
 Best Local Similarity 46.6%; Pred. No. 9.2e-41;
 Matches 95; Conservative 44; Mismatches 61; Indels 4; Gaps 2;
 QY 1 MAGKSLFKVILLGGGVGKSSLMRYVTNKFDTQLFHTIGVEFLNKDLEVDGHFVTMQI 60
 DB 1 MPSPANVLLVILGDSGVGKTSMLNQYVNRKFSNQYKATIGADFLTKVEQFEDRLFTLQI 60
 QY 61 WTAGQRRFSLRTPFYRGSDCLLTFFVDDSQSFQNLNWKKEFIYYADVKEPSPFV 120
 DB 61 WTAGQRRFSLRTPFYRGSDCLLTFFVDDSQSFQNLNWKKEFIYYADVKEPSPFV 120
 QY 121 ILGNKIDI---SERQVSTEEAQAQWCRDNGDYYPFETSAKDATNVAAAFEEAVRVLAT 177
 DB 121 VLGNKVIDDGNRSRVSEKKVAKWASKGNIPYFETSAKEGTFVEEAFQCIKNAKSGE 180
 QY 178 RSDHLIQTDTVNLHRKPKPSSSCC 201
 DB 181 -EEIYLPDTLDVGVSSQRTGCG 203

RESULT 7
 Q40526 PRELIMINARY; PRT; 206 AA.
 ID Q40526
 AC Q40526
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Nt-rab7a protein.
 GN NT-RAB7A.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; lamids; Solanales; Solanaceae; Nicotiana.
 OX NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SRI;
 RX MEDLINE=95303981; PubMed=7784525;
 RA Haizel T., Merkle T., Turck F., Nagy F.;
 RT "Characterization of membrane-bound small GTP-binding proteins from
 RT Nicotiana tabacum."
 RL Plant Physiol. 108:59-67(1995).
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; L29274; AAA74118.1; -.
 DR HSSP; P36017; 1EK0.

DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 206 AA; 23104 MW; 59CDBCDBDA4794294 CRC64;
 Query Match 48.4%; Score 514; DB 10; Length 206;
 Best Local Similarity 47.8%; Pred. No. 1.2e-40;
 Matches 97; Conservative 40; Mismatches 62; Indels 4; Gaps 2;
 QY 2 AGKSLFKVILLGGGVGKSSLMRYVTNKFDTQLFHTIGVEFLNKDLEVDGHFVTMQI 61
 DB 3 ARRRMLKVILGDSGVGKTSMLNQYVNRKFSNQYKATIGADFLTKVEQFEDRLFTLQIW 62
 QY 62 DTAGQRRFSLRTPFYRGSDCLLTFFVDDSQSFQNLNWKKEFIYYADVKEPSPFVI 121
 DB 63 DTAGQRRFSLRTPFYRGSDCLLTFFVDDSQSFQNLNWKKEFIYYADVKEPSPFVI 122
 QY 122 LGNKIDI---SERQVSTEEAQAQWCRDNGDYYPFETSAKDATNVAAAFEEAVRVLAT 178
 DB 123 LGNKIDVDGNSRVSEKKVAKWASKGNIPYFETSAKEGTFVEEAFQCIKNAKNEP- 181
 QY 179 SDHLIQTDTVNLHRKPKPSSSCC 201
 DB 182 EDEIYLPDTLDVAGSQSRSTGC 204
 RESULT 8
 Q8W0B7 PRELIMINARY; PRT; 217 AA.
 ID Q8W0B7
 AC Q8W0B7
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE RAS-related GTP-binding protein Rab7 family.
 GN P0452F10.6.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone:P0452F10."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 DR EMBL; AP003434; BAB78669.1; -.
 DR HSSP; P05713; 3RAB.
 DR Gramene; Q8W0B7; -.
 DR InterPro; IPR003579; GTPase_Rab.
 DR InterPro; IPR001806; Ras_trnsmg.
 DR InterPro; IPR002078; Sig54_interact.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF00071; ras; 1.
 DR PRINTS; PR00449; RASTNSFRMG.
 DR SMART; SM00175; RAB; 1.
 DR TIGRFAMS; TIGR00231; small_GTP; 1.
 DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
 KW GTP-binding; Lipoprotein.
 SQ SEQUENCE 217 AA; 24361 MW; 7F32441A2F08183B CRC64;
 Query Match 48.4%; Score 513.5; DB 10; Length 217;
 Best Local Similarity 52.5%; Pred. No. 1.4e-40;
 Matches 96; Conservative 34; Mismatches 48; Indels 5; Gaps 2;

```

QY 7 LFKVILLGGGVGKSSLMNRVYTNKPTQLFHTIGVEFLNKLDEVDGHFVTMOIWDTAGQ 66
DB 8 LLKVILLGDSGVGKISLMNQVNVKFNQYKATIGADFLLTKEVQIDRLLQLQIWDTAGQ 67
QY 67 ERPSLTPFPYRGSDCCLLTFSVDDSGSQNLNWKKEFYIYADVKEPSPFPFVILGNKI 126
DB 68 ERFQSLGVAFYRGADCCVLVDVNVTKSFERLNSWREFFLIQASPSDPENFPFVILGNKI 127
QY 127 DI---SRRQVSTEAQACRDNGDYPYFETSADATNVAAPFEAVRRVLATEDRSD--H 181
DB 128 DVDGNGERTSEKKAKAWCASKGNIPYFETSAKEGFVNEAFCIARNAIKNPESEIIF 187
QY 182 LIQ 184
DB 188 LIQ 190

RESULT 9
Q40213 PRELIMINARY; PRT; 206 AA.
AC Q40213
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RAB7C.
GN RAB7C.
OS Lotus japonicus.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae; Lotus.
OX NCBI_TaxID=34305;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Root nodules; PubMed=9076991;
RX MEDLINE=97231679; PubMed=9076991;
RA Borg S., Brandstrup B., Jensen T.J., Poulsen C.;
RT "Identification of new protein species among 33 different small GTP-
RT binding proteins encoded by cDNAs from Lotus japonicus, and expression
RT of corresponding mRNAs in developing root nodules.";
RL Plant J. 11:237-250(1997).
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; Z73942; CAA98170.1; -.
DR HSSP; P36017; 1EKO.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.
DR SMART; SM00175; RAB; 1.
DR TIGFAMS; TIGR00231; small GTP; 1.
DR PROSITE; PS00675; SIGMAS4_INTERACT_1; 1.
KW GTP-binding; Lipoprotein.
SQ SEQUENCE 206 AA; 23025 MW; F2453D6B1705177F CRC64;

Query Match
Best Local Similarity 48.3%; Score 513; DB 10; Length 206;
Matches 96; Conservative 40; Mismatches 61; Indels 4; Gaps 2;

QY 4 KSLFKVILLGGGVGKSSLMNRVYTNKPTQLFHTIGVEFLNKLDEVDGHFVTMOIWDWT 63
DB 5 RRTLLKVIILGDSGVGKISLMNQVNVKFNQYKATIGADFLLTKEVQIDRLLQLQIWDWT 64
QY 64 AQGERFRSLTPFYRGSDCCLLTFSVDDSGSQNLNWKKEFYIYADVKEPSPFPFVILG 123
DB 65 AQGERFQSLGVAFYRGADCCVLVDVNVTKSFERLNSWREFFLIQASPSDPENFPFVILG 124
QY 124 NKIDI---SERQVSTEAQACRDNGDYPYFETSADATNVAAPFEAVRRVLATEDRSD 180
DB 125 NKIDIDGNGSRVSVSEKKARAWCASKGNIPYFETSAKGGINVEEAFQCIAKNALKSGE-EE 183
QY 181 HLIQDTVNLNHRKPKPSSCC 201

```

```

DB 184 ELYLPDITDVGNSSQORASGC 204

RESULT 10
Q9LS94 PRELIMINARY; PRT; 206 AA.
AC Q9LS94
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE RAS-related GTP-binding protein (Atrab71) (Putative RAS-related GTP-
DE binding protein).
GN ATRAB71 OR MVE11.21/AT3G18820.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Columbia;
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Columbia;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC
RT clones.";
RL DNA Res. 7:131-135(2000).
RN [3]
RP SEQUENCE FROM N.A.
RX Ueda T., Wada Y., Nakano A.;
RT "Rab7 homologs in Arabidopsis thaliana.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,
RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Koesema E.,
RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,
RA Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A.,
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MVE11.21/AT3G18820 (GI:9293907)";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
RA Shinn P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,
RA Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MVE11.21/AT3G18820 (GI:9293907)";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AB026654; BAB01810.1; -.
DR EMBL; AB071846; BAB68371.1; -.
DR EMBL; AY035137; AAK59641.1; -.
DR EMBL; AY059072; AAL15178.1; -.
DR HSSP; P36017; 1EKO.
DR InterPro; IPR003579; GTPase_Rab.
DR InterPro; IPR001806; Ras_trnsfrmng.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMNG.

```


RESULT 12
 Q9VIM6 PRELIMINARY; PRT; 256 AA.
 AC Q9VIM6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE CG9994 protein (R017845p).
 GN CG9994.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K.J., Evans G., Fouts D.E., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foaier C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Sidgen-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 SC Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dreenek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guatin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunco J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celnik S.,
 RA Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RL -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
 CC EMBL; AB003662; AAF53798.1; --
 DR EMBL; AY071139; AAL48761.1; --
 DR HSP; P05713; 3RAB.
 DR FlyBase; Fgn0032782; CG9994.
 QY 4 KSLFKVILLGDDGGVKGKSLMNRVYTNKFTQFHTIGVEFLNKDLVDGHFVTMQIWDT 63

```

Db 5 KVLKVLIIIDSGVSGKSLMNQVNVKFSYATIGADFLFREVLVDQRVTQMLWDT 64
Qy 64 AGQRFSLRTPFYRGSDCLLTFSDVDSQFQNLNWKKEFYIYADVKEPESFPFVILG 123
Db 65 AGQRFQSLGVAFYRGADCCVLYDVNNSKFDALDSWRDEFLIQASPRDPNFPFVILG 124
Qy 124 NKIDI--SERQVSTEEACACWCRNGDYPYETSNGATNVAAPFAEAVRVLATERSDH 181
Db 125 NKIDMESKRVISYTRMTWCQSGKNIPYFETSAKEAINVEQAFVIARNALMQEESBF 184
Qy 182 LIQ-TDVTNLRHKPKPSSSCC 201
Db 185 SGDFQDDPNIHENDRGAC 205

RESULT 14
Q9C820 PRELIMINARY; PRT; 206 AA.
AC Q9C820;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GTP-binding protein RAB7D, putative (Acrab72) (Putative GTP-binding protein RAB7D).
GN F19K6.10 OR ATRAB72 OR ATIG52280.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucots II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S., White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y., Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W., Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K., Dunn P., Etgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y., Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L., Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E., Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B., Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P., Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A., Millscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I., Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M., Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D., Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M., Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RA "Sequence and analysis of chromosome 1 of the plant Arabidopsis thaliana."
RL Nature 408:816-820 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Ueda T., Wada Y., Nakano A.;
RT "Rab7 homologs in Arabidopsis thaliana."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M., Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RA "Full Length cDNA of gene AtIG52280 (GI:15218194).";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.

```

```

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M., Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RA "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC037424; AAG51552.1; -
DR EMBL; AB071847; BAB68372.1; -
DR EMBL; AY074361; AAL67057.1; -
DR EMBL; AY096407; AAM20047.1; -
DR HSP; F36017; IEXO.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR TIGSPAMs; TIGR00231; small_GTP; 1.
DR PROSITE; PS00675; SIGMA54_INTERACT_1; 1.
KW GTP-binding.
SQ SEQUENCE 206 AA; 23069 MW; 6E31BB86CD75CB7F CRC64;
Query Match 47.9%; Score 509; DB 10; Length 206;
Best Local Similarity 48.5%; Pred. No. 3.4e-40;
Matches 96; Conservative 37; Mismatches 61; Indels 4; Gaps 2;
Qy 7 LFKVILLGDGVGKSLMNRYVTKPDTQLFHTTIGVEFLNKLVDGHFVTQIWDTAGQ 66
Db 8 LLKVIILLGDSGVGKSLMNQVNVKFSYATIGADFLFREVLVDQRVTQMLWDTAGQ 67
Qy 67 ERFSLRTPFYRGSDCLLTFSDVDSQFQNLNWKKEFYIYADVKEPESFPFVILGKI 126
Db 68 ERFQSLGVAFYRGADCCVLYDVNNSKFDALDSWRDEFLIQASPRDPNFPFVILGKI 127
Qy 127 DI---SERQVSTEEACACWCRNGDYPYETSNGATNVAAPFAEAVRVLATERSDHLI 183
Db 128 DVDGKGRVSEKKAKAWCAKSKGNIPYFETSAKEGFGNVDAFAFCITKNAFKNPEEPY 187
Qy 184 QTDVTNLRHKPKPSSSCC 201
Db 188 -PDTIDVAGGQQQRSTGC 204

RESULT 15
Q9NHWS PRELIMINARY; PRT; 206 AA.
AC Q9NHWS;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Rab7-like GTPase (Fragment).
GN RAB7.
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RA Temesvari L.A.;
RT "A Rab7-like GTPase of Entamoeba histolytica."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO RAS PROTEINS. BELONGS TO THE RAB SUBFAMILY.
DR EMBL; AF218311; AAF32317.1; -
DR HSP; F36017; IEXO.
DR InterPro; IPR003579; GTPase Rab.
DR InterPro; IPR001806; Ras_trnsfrmg.
DR InterPro; IPR002078; Sig54_interact.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF00071; ras; 1.
DR PRINTS; PR00449; RASTRNSFRMG.
DR SMART; SM00175; RAB; 1.

```

Search completed: February 10, 2004, 04:58:12
Job time : 90 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 23:58:36 ; Search time 4598 Seconds
(without alignments)
10477.073 Million cell updates/sec

Title: US-09-988-974-6
Perfect score: 1175
Sequence: 1 CACTGTGATGAACACTTTT.....GTAAAAA.....AAAAA 1175

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rnd.*
- 36: em_htg_rnd.*
- 37: em_htg_vit.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1175	100.0	1175	6	AR034083	AR034083 Sequence
2	1100	93.6	213863	9	AC003037	BC017265 Homo sapi
3	993.2	84.5	1106	9	BC017265	BC017265 Homo sapi
C 4	818.2	69.6	143872	9	AC091987	AC091987 Homo sapi
5	713.6	60.7	1001	4	CFRAB9	X58386 Canine rab9
C 6	696.6	59.3	130407	2	AC084354	AC084354 Homo sapi
7	696.6	59.3	156357	9	AC022486	AC022486 Homo sapi
C 8	686.2	58.4	183293	9	AC142348	AC142348 Pan trogl
C 9	682.8	58.1	161879	9	AC017005	AC017005 Homo sapi
10	682.8	58.1	166436	9	AC016752	AC016752 Homo sapi
C 11	682.8	58.1	182083	9	AC007965	AC007965 Homo sapi
12	664.6	56.6	1251	10	BC008160	BC008160 Mus muscu
13	660.2	56.2	1276	10	AB027290	AB027290 Mus muscu
C 14	656.6	55.9	201642	10	AL672174	AL672174 Mouse DNA
15	646.6	55.0	184474	10	AL672047	AL672047 Mouse DNA
16	646.6	55.0	204031	2	AC093354	AC093354 Mus muscu
17	638.2	54.3	208391	2	AC130009	AC130009 Rattus no
C 18	638.2	54.3	217558	2	AC111489	AC111489 Rattus no
19	632.4	53.8	1286	10	AF325692	AF325692 Rattus no
20	604.4	51.4	606	6	AR270519	AR270519 Sequence
21	604.4	51.4	606	9	AF498944	AF498944 Homo sapi
22	604.4	51.4	606	9	HSU44103	U44103 Human small
23	492	41.9	64450	9	AC073893	AC073893 Homo sapi
C 24	492	41.9	145383	9	AC010682	AC010682 Homo sapi
25	492	41.9	149794	2	AC007273	AC007273 Homo sapi
C 26	439.6	37.4	487	6	AR237295	AR237295 Sequence
C 27	439.6	37.4	487	6	AR260788	AR260788 Sequence
C 28	439.6	37.4	487	6	AR278319	AR278319 Sequence
C 29	439.6	37.4	487	6	AX106425	AX106425 Sequence
C 30	439.6	37.4	487	6	AX140716	AX140716 Sequence
C 31	439.6	37.4	487	6	AX200576	AX200576 Sequence
C 32	439.6	37.4	487	6	AX267232	AX267232 Sequence
C 33	439.6	37.4	487	6	BD070348	BD070348 Compounds
34	426.4	36.3	623	9	HSU44105	U44105 Human Rab9
35	385.4	32.8	399	6	BD054210	BD054210 Sequence
36	333.2	28.4	1049	9	AB036693	AB036693 Homo sapi
C 37	333.2	28.4	31483	9	AL139228	AL139228 Human DNA
C 38	333.2	28.4	171430	2	AC027160	AC027160 Homo sapi
C 39	333	28.3	199658	2	AC133053	AC133053 Rattus no
C 40	329.2	28.0	212953	2	AC144376	AC144376 Pan trogl
C 41	328.4	27.9	160571	10	AL672008	AL672008 Mouse DNA
C 42	328.2	27.9	163495	9	AC024183	AC024183 Homo sapi
C 43	328.2	27.9	164891	2	AC022848	AC022848 Homo sapi
44	328.2	27.9	199174	9	AC007742	AC007742 Homo sapi
C 45	324.8	27.6	158422	9	AC026616	AC026616 Homo sapi

ALIGNMENTS

RESULT 1
AR034083
LOCUS AR034083 1175 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 6 from patent US 5869291.
ACCESSION AR034083
VERSION AR034083.1 GI:5949688
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1175)
AUTHORS Hillman,J.L., Goli,S.K. and Bandman,O.
TITLE Rab proteins
JOURNAL Patent: US 5869291-A 6 09-FEB-1999;
FEATURES Location/Qualifiers

ambiguities. If the sequence quality does not meet this standard, it will be indicated in the annotation.

The repeat regions shown were identified using RepeatMasker by Adrian Smit.

Sequence similarities were identified using Powerblast by Jinghui Zhang.

Exon/intron boundaries of identified genes were chosen if there were canonical splice junctions that maintained sequence continuity across the splice junctions.

FEATURES

source

```
1..213863
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/map="Xp22"
/clone="GSHB-526D21"
/clone_lib="Genome Systems Human BAC Library"
```

```
1..1999
/note="Region: AA029104 zk09b01.r1 Soares pregnant uterus complement(380..697)"
/function="Overlaps with adjacent clone AC079383"
```

```
1..303
/note="Region: AA043674 zk50a09.s1 Soares pregnant uterus complement(1102..1291)"
/rpt_family="POLY A"
```

```
1..1387
/rpt_family="MER3"
```

```
1..2143
/note="Region: AA740623 ny97g07.s1 NCI CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1286268"
```

```
complement(2794..2899)
/rpt_family="L2"
```

```
3015..3044
/rpt_family="AT-rich"
```

```
complement(3128..3423)
/rpt_family="AluJb"
```

```
complement(3949..4245)
/rpt_family="AluJb"
```

```
complement(4247..4607)
/rpt_family="L1MB7"
```

```
6043..6085
/rpt_family="L2"
```

```
6389..6517
/rpt_family="MIR"
```

```
complement(6975..7023)
/rpt_family="L2"
```

```
complement(7007..7092)
/rpt_family="MIR"
```

```
complement(7787..8093)
/rpt_family="AluSx"
```

```
complement(8291..8318)
/rpt_family="AT-rich"
```

```
complement(9569..9872)
/rpt_family="AluSx"
```

```
complement(10250..10411)
/rpt_family="MER3"
```

```
11214..11510
/rpt_family="AluJc"
```

```
11602..11660
/rpt_family="MIR"
```

```
complement(12431..12991)
/note="Region: AA563671 ng49e02.s1 NCI CGAP_C03 Homo sapiens cDNA clone IMAGE:1938138 similar to 5W:RAB9_CANPA
```

```
P24408 RAS-RELATED PROTEIN RAB-9"
```

```
13871..14137
/rpt_family="AluSc"
```

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```
15421..15925
/note="Region: AA828182 of33g12.s1 NCI CGAP_Kid6 Homo sapiens cDNA clone IMAGE:1422982"
complement(16053..16074)
/rpt_family="AT-rich"
16136..16293
/rpt_family="AluSp"
16342..16644
/rpt_family="AluSx"
16710..16999
/rpt_family="AluSg"
17038..17335
/rpt_family="AluSc"
18365..18654
/rpt_family="AluSx"
20209..20243
/rpt_family="POLY A"
20634..20920
/rpt_family="AluSx"
20946..21375
/rpt_family="MLT1C"
22045..22349
/rpt_family="AluSx"
23918..24220
/rpt_family="AluJb"
24370..24596
/rpt_family="MIR"
complement(24994..25130)
/rpt_family="MER58A"
25383..25414
/rpt_family="POLY A"
25856..25953
/rpt_family="MER3"
25956..26288
/rpt_family="AluY"
complement(26289..26373)
/rpt_family="MER33"
complement(26397..26641)
/rpt_family="AluSx"
26965..27292
/rpt_family="MLT1D"
complement(27563..27593)
/rpt_family="AT-rich"
27807..32800
/rpt_family="L1PA2"
32380..32557
/standard_name="STS1-CSRL-31b6-ub/CSRL-31b6-uz, Chr. X, Homo sapiens"
Homo sapiens
/db_xref="dbSTS:7821"
32801..33102
/rpt_family="AluSp"
33106..34245
/rpt_family="L1PA2"
34748..34864
/rpt_family="L2"
35687..35752
/standard_name="HUMSWX1791, Chr. X, Homo sapiens"
/db_xref="dbSTS:13087"
37266..37472
/rpt_family="MER20"
complement(40152..40194)
/rpt_family="L2"
complement(40214..40291)
/rpt_family="MIR"
complement(40560..40761)
/rpt_family="L1MC1"
40763..40844
/rpt_family="AluSg/x"
complement(41023..41322)
/rpt_family="AluSx"
complement(42296..42322)
/rpt_family="AT-rich"
42638..42936
```

Query Match		93.6%	Score 1100;	DB 9;	Length 213863;
Best Local Similarity		100.0%	Pred. No. 9.3e-209;		
Matches 1100;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	59	AGGGTCTTGAAGCTTTTGAGATTAAACATGGCAGGAAATCATCTTTTAAAGTAAT	118		
Db	11893	AGGGTCTTGAAGCTTTTGAGATTAAACATGGCAGGAAATCATCTTTTAAAGTAAT	11952		
Qy	119	TCCTCTGGAGATGGTGGAGTTGGGAGAGTTCACTTATGAACAGATATGTAACTAATA	178		
Db	11953	TCCTCTGGAGATGGTGGAGTTGGGAGAGTTCACTTATGAACAGATATGTAACTAATA	12012		
Qy	179	GTTCGATACCCAGCTTTCATCAATAGGTGTGGAATTTTAAATAAAGATTTGGAAGT	238		
Db	12013	GTTCGATACCCAGCTTTCATCAATAGGTGTGGAATTTTAAATAAAGATTTGGAAGT	12072		
Qy	239	GGATGACATTTTGTACATCCAGATTTGGGACACGGCAGGTTCAGGACGATTCGGAAG	298		
Db	12073	GGATGACATTTTGTACATCCAGATTTGGGACACGGCAGGTTCAGGACGATTCGGAAG	12132		
Qy	299	CCTGAGGACACCATTTTACAGAGGTTCTCACTGCTGCTTACTTTAGTGTGATGA	358		
Db	12133	CCTGAGGACACCATTTTACAGAGGTTCTCACTGCTGCTTACTTTAGTGTGATGA	12192		
Qy	359	TTCAAAAGCTTCAGAACTTAAGTAACTGGGAAGAAAGAAATTCATATATTGCGATGT	418		
Db	12193	TTCAAAAGCTTCAGAACTTAAGTAACTGGGAAGAAAGAAATTCATATATTGCGATGT	12252		
Qy	419	GAAGAGCCTGAGAGCTTTCCTTTTGTGATTTCTGGTTAAACAGATTTGACATAAGCGAAG	478		
Db	12253	GAAGAGCCTGAGAGCTTTCCTTTTGTGATTTCTGGTTAAACAGATTTGACATAAGCGAAG	12312		
Qy	479	CGAGTGTCTACAGAAAGCCCAAGCTTGGTGACGGCAACCGCGATATCTTATTT	538		
Db	12313	CGAGTGTCTACAGAAAGCCCAAGCTTGGTGACGGCAACCGCGATATCTTATTT	12372		
Qy	539	TGAACAGTGCAGAGTGCACAAATGTGGCAGCAGCTTTGAGAGCGGTTTCCGAG	598		
Db	12373	TGAACAGTGCAGAGTGCACAAATGTGGCAGCAGCTTTGAGAGCGGTTTCCGAG	12432		
Qy	599	AGTCTTGTCTACGAGGATAGTTCAGATCATTTGATTCAGACAGACAGTCAATCTTCA	658		
Db	12433	AGTCTTGTCTACGAGGATAGTTCAGATCATTTGATTCAGACAGACAGTCAATCTTCA	12492		
Qy	659	CGAAAGCCCAAGCCTAGCTCATCTTGTCTGTTGATTTGATTTGATTTGATTTGATTT	718		
Db	12493	CGAAAGCCCAAGCCTAGCTCATCTTGTCTGTTGATTTGATTTGATTTGATTTGATTT	12552		
Qy	719	CAACTCACATATACACAAATCAACATGGGATGGAGAGAGATTTAGCGTTTGCAGC	778		
Db	12553	CAACTCACATATACACAAATCAACATGGGATGGAGAGAGATTTAGCGTTTGCAGC	12612		
Qy	779	AGTGTATCATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	838		
Db	12613	AGTGTATCATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	12672		
Qy	839	GGGACACATCCCTCTTGGAGGAATATATTTACTCAATAATGGCACCTTTACATTTATAA	898		
Db	12673	GGGACACATCCCTCTTGGAGGAATATATTTACTCAATAATGGCACCTTTACATTTATAA	12732		
Qy	899	TTGTAACAGTGTCTAATAACGTTCTTTTAAATTTAAATGTAAGTTCAGAGCTAATAA	958		
Db	12733	TTGTAACAGTGTCTAATAACGTTCTTTTAAATTTAAATGTAAGTTCAGAGCTAATAA	12792		
Qy	959	ATGAATGACCAAGACTTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA	1018		
Db	12793	ATGAATGACCAAGACTTTTAAATTAATAAATAAATAAATAAATAAATAAATAAATAA	12852		
Qy	1019	TACTTGGATTTTCTCGGAAATAGGAACTACTTTTTATATGATGTATGTTTTATGC	1078		
Db	12853	TACTTGGATTTTCTCGGAAATAGGAACTACTTTTTATATGATGTATGTTTTATGC	12912		
Qy	1079	AATTAGCATTTCTTCTTGGTTTCAGGGAATACCTTCTCTAAGCAATAATCTTAGATATT	1138		
Db	12913	AATTAGCATTTCTTCTTGGTTTCAGGGAATACCTTCTCTAAGCAATAATCTTAGATATT	12972		
Qy	1139	AAAGATTAAATCTAATGTA	1158		
Db	12973	AAAGATTAAATCTAATGTA	12992		
RESULT 3					
BC017265					
LOCUS					
DEFINITION	Homo sapiens, RAB9, member RAS oncogene family, Clone MGC:29629				
ACCESSION	BC017265				
VERSION	BC017265.1				
KEYWORDS	MGC.				
SOURCE	MGC.				
ORGANISM	Homo sapiens (human)				
REFERENCE	1 (bases 1 to 1106)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (05-NOV-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk				
	Email: cgaps-xemail.nih.gov				
	Tissue Procurement: ATCC				
	cDNA Library Preparation: Rubin Laboratory				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)				
	DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;				
	Web site: http://www.nisc.nih.gov/				
	Contact: nisc.mgc@nhgri.nih.gov				
	Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantipop, S., Thomas, P.J., Tongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.				
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov					
Series: IRAL Plate: 39 Row: c Column: 18.					
Location/Qualifiers					
1. 1106					
/organism="Homo sapiens"					
/mol_type="mRNA"					
/db_xref="LocusID:9367"					
/db_xref="taxon:9606"					
/clone="MGC:29629 IMAGE:4139714"					
/tissue_type="Muscle, rhabdomyosarcoma"					
/clone_lib="NIH MGC 17"					
/lab_host="DH10B-R"					
/note="Vector: pOTB7"					
192. 797					
/codon_start=1					
/product="RAB9, member RAS oncogene family"					
/protein_id="AAH17265.1"					
/db_xref="GI:16878118"					
/translation="MAGKSLFKVILLGGVGKSLMRYVTKNFKDTQLFHTIGVEF LNKLEVDGHFTVQIWDVAGRFRSLTFPFGSDCCLLTFVDDSGSFQNLNWK KFFIYADVKEPESFPVILGNKIDISERQVSTEEAQAACWRDNGDYPYFETSAKDATN VAAAEAVRRVLATEDRSDHLIDTIVNLHKKPSSSCC"					
BASE COUNT 339 a 210 c 249 g 308 t					
ORIGIN					

Query Match		84.5%;	Score 993.2;	DB 9;	Length 1106;
Best Local Similarity		99.7%;	Pred. No. 4.4e-187;		
Matches 995;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	CACCTGTGATGAAACACTTTTCCCGTGTGCTTTGAGTGCATCTTCTCAACACCCCTAGGAG	60		
Db	106	CACCTGTGATGAAACACTTTTCCCGTGTGCTTTGAGTGCATCTTCTCAACACCCCTAGGAG	165		
QY	61	GCTTCTTCAAGCTTTTGGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC	120		
Db	166	GCTTCTTCAAGCTTTTGGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC	225		
QY	121	TCCTTGGAGATGGTGGAGCTTGGAGAGCTTCACTTATGAAACAGATATGTAACATAAGT	180		
Db	226	TCCTTGGAGATGGTGGAGCTTGGAGAGCTTCACTTATGAAACAGATATGTAACATAAGT	285		
QY	191	TTGATACCCAGCTCTTCCATACATAGTGTGGAAATTTTAAATAAAGATTGGAAGTGG	240		
Db	286	TTGATACCCAGCTCTTCCATACATAGTGTGGAAATTTTAAATAAAGATTGGAAGTGG	345		
QY	241	ATGGACATTTTGTACCATGAGATTGGGACACGGCAGGTCCAGAGCGATTCCGAAGCC	300		
Db	346	ATGGACATTTTGTACCATGAGATTGGGACACGGCAGGTCCAGAGCGATTCCGAAGCC	405		
QY	301	TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTGATGATT	360		
Db	406	TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTGATGATT	465		
QY	361	CACAAAGCTTCCAGAACTTAACTAGTAACCTGGAAGAAATTCATATATTATGAGATGTA	420		
Db	466	CACAAAGCTTCCAGAACTTAACTAGTAACCTGGAAGAAATTCATATATTATGAGATGTA	525		
QY	421	AAGAGCTGAGAGCTTTTCTTTTGTGATTTCTGGGTAAACAAGTTGACATAGCGAACCGC	480		
Db	526	AAGAGCTGAGAGCTTTTCTTTTGTGATTTCTGGGTAAACAAGTTGACATAGCGAACCGC	585		
QY	481	AGGTGTCTACAGAAAGACCCAGCTTGTGTCAGGACCAACGGCGACTTCCTTATTTCG	540		
Db	586	AGGTGTCTACAGAAAGACCCAGCTTGTGTCAGGACCAACGGCGACTTCCTTATTTCG	645		
QY	541	AAACAAGTGCAAAAGATGCCAACAAATGTGCAGCAGCCTTTGAGGAAGCGGTTTGAAGAG	600		
Db	646	AAACAAGTGCAAAAGATGCCAACAAATGTGCAGCAGCCTTTGAGGAAGCGGTTTGAAGAG	705		
QY	601	TTCTTGCTACCGAGTAGGTGCAGATCATTTGATTCAGACAGACAGATCAATCTTCACC	660		
Db	706	TTCTTGCTACCGAGTAGGTGCAGATCATTTGATTCAGACAGACAGATCAATCTTCACC	765		
QY	661	GAAGGCCAAGCCTAGCTCATCTTGTGTTGATTTGTAGATTGTTGATGCAATCTTAACCA	720		
Db	766	GAAGGCCAAGCCTAGCTCATCTTGTGTTGATTTGTAGATTGTTGATGCAATCTTAACCA	825		
QY	721	ACTCAGACATATACAGAAATCAACATGGGGATGGGAGAGAGAAATTAGCGTTTGCAGCAG	780		
Db	826	ACTCAGACATATACAGAAATCAACATGGGGATGGGAGAGAGAAATTAGCGTTTGCAGCAG	885		
QY	781	TGTATCATCTPACTAATAAATTAACCTAATGTTGCTGCTTCAATTAGTTGTTGGGAGAGG	840		
Db	886	TGTATCATCTPACTAATAAATTAACCTAATGTTGCTGCTTCAATTAGTTGTTGGGAGAGG	945		
QY	841	GACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGACCTTCAATTTATAAATT	900		
Db	946	GACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGACCTTCAATTTATAAATT	1005		
QY	901	GTAACAGTTGTCTAATAACCTTTCTTTTAAATTAATATGATGTTGACAGCTAATAAT	960		
Db	1006	GTAACAGTTGTCTAATAACCTTTCTTTTAAATTAATATGATGTTGACAGCTAATAAT	1065		
QY	961	GAATGACCAAGACTTAAATTAATAATAAATAAGAAA	998		
Db	1066	GAATGACCAAGACTTAAATTAATAATAAATAAATAA	1103		
RESULT 4					
AC091987/c					
LOCUS					
DEFINITION					
AC091987					
AC091987					
AC091987.2					
GT:20376969					
HTG.					
Homo sapiens (human)					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
1 (bases 1 to 143872)					
DOE Joint Genome Institute and Stanford Human Genome Center.					
Direct Submission					
Unpublished					
2 (bases 1 to 143872)					
DOE Joint Genome Institute.					
Direct Submission					
Submitted (09-JUN-2001)					
Production Sequencing Facility, DOE Joint					
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA					
3 (bases 1 to 143872)					
DOE Joint Genome Institute and Stanford Human Genome Center.					
Direct Submission					
Submitted (01-MAY-2002)					
DOE Joint Genome Institute, 2800 Mitchell					
Drive, Walnut Creek, CA 94598, USA					
On May 1, 2002 this sequence version replaced gi:14333923.					
Draft Sequence Produced by DOE Joint Genome Institute					
www.jgi.doe.gov					
Finishing Completed at Stanford Human Genome Center					
www.shgc.stanford.edu					
Quality: Phrap Quality >=40 99.4% of Sequence;					
Estimated Total Number of Errors is 0.6.					
NOTE: This insert is not the entire sequence of the clone (entire					
sequence is 149.5kb). It is clipped at the overlap with AC010590.					
The number of bases overlapped is 20375.					
FEATURES					
location/Qualifiers					
1..143872					
/organism="Homo sapiens"					
/mol_type="genomic DNA"					
/db_xref="taxon:9606"					
/chromosome="5"					
/clone="RP11-8E2"					
BASE COUNT					
ORIGIN					
45872 a 25416 c 25617 g 46967 t					
Query Match					
Best Local Similarity					
Matches 970;					
Conservative 0;					
Mismatches 113;					
Indels 37;					
Gaps 5;					
QY	86	AATGCGCAGGAAATCATCACTTTTAAAGTAATTCCTCTGGAGATGGTGGAGTTGGAA	145		
Db	26969	AACGCGAGGAAATCATCACTTTTAAAGTAATTCCTCTGGAGATGGTGGAGTTGGAA	26910		
QY	146	GAGTTCACTTATGAACAGATATGTAACCTAATAGTTTGATACCCAGCTTCCATCAAT	205		
Db	26909	GAGTTCACTTATGAACAGATATGTAACCTAATAGTTTGATACCCAGCTTCCATCAAT	26850		
QY	206	AGGTGTGGAATTTTAAATTAAGATTGGAAGTGGATGGATTTGGTACCAAGTAGAA	265		
Db	26849	AGGTGTGGAATTTTAAATTAAGATTGGAAGTGGATGGATTTGGTACCAAGTAGAA	26790		
QY	266	TTGGGACACGCGCAGTCCAGGAGGATTCGAGCTGAGACACCATTTTACAGAGTTTC	325		
Db	26789	TTGGGACACGCGCAGTCCAGGAGGATTCGAGCTGAGACACCATTTTACAGAGTTTC	26730		
QY	326	TGACTGCTGCTGCTTACTTTTGTAGTTCGATGTTTCAAAAGCTTCCAGAACTTAAGTAA	385		
Db	26729	TGACTGCTGCTGCTTACTTTTGTAGTTCGATGTTTCAAAAGCTTCCAGAACTTAAGTAA	26671		
QY	386	CTGGAACAGAGAAATTCATATATTATGAGATGTAAGAGCCCTGAGAGCTTCTCTTTGT	445		
Db	26670	CTGGAACAGAGAAATTCATATATTATGAGATGTAAGAGCCCTGAGAGCTTCTCTTTGT	26611		

```
Qy 446 GATTCTGGTAAACAAGATTGACATAGCGAAACGCGAGGTGTCTACAGAGAAGCCCAAGC 505
Db 26610 AATTTTGGGTAAACAAGCTGAC---AGTGAACAGCAGGTGTCTGCAAGAGAAGCCCAAGC 26554
Qy 506 TTGGTGAGGAGCAACGGCGACTATCTCTTATTTTGAACAAGATGCAAAAGATGCCACAAA 565
Db 26553 CTGGTGAGGAGGAGCGAGCAGCTATCTCTTAAATTTGAACAAGATGCAAAAGATGCCACAAA 26494
Qy 566 TTGGGAGCAGCCTTTTGAAGAGCGGTTGGAAGAGTTCTTGCTACCGAGATAGTGTGAGA 625
Db 26493 TTGGGAGATAGCAATTTGAAGAGCGGTTCCAGAGTTCTTACTACCGAGATAGTGTGAGA 26434
Qy 626 TCATTGATTTCAGACAGACA-----CACTCAATCTTCCACCA 662
Db 26433 TCATTGACTCAGAGAGTGATGCTGCTGACTTCATCGAAACCACTTGTCTCAATCTTCATTGA 26374
Qy 663 AAGCCCAAGCCTAGCTCATCTGCTGCTGATGTT-----AGATTGTTGATGATCTTAAC 718
Db 26373 AAGCCCAAGCCTAGCTCATCTTAAATGTTGATGTTTAAAGACATTTGTTGATGATCTTAA 26314
Qy 719 CAACTCACACATATACACAAAATCAACATGGGATGGAGAAGAGATTAGCGTTTGCAGC 778
Db 26313 CAACTCACAGGTACACACAAAATCAACATGGGATGGAGAAGAGATTAGCATTTGCAGC 26254
Qy 779 AGTGATCATCTACTAATAAATAAATAAATAATGT-----TGCTGCTTCAATAGTTGGTG 832
Db 26253 AGTGATCATCTACTAATAAATAAATAAATAATGTATAGTGTGCTTTTCAATAGTTGGTA 26194
Qy 833 GGAGAAGGAGCACATCCACTCTTGGAGGAATATATTTACTCAATATGACCACTTACATT 892
Db 26193 GGAGAAGGAGCACATCCACTCATGGAGGAGTCTGTTTACTGATATGACCACTTACATT 26134
Qy 893 TATAAATTGTAAGTGTCTAATAAGTTCCTTAATTAATTAATTAATTAATTAATTAATTA 952
Db 26133 TATAAATTCGAAGTGTGTTTAAACACATTTCTTAAATTAATTAATTAATTAATTAATTA 26074
Qy 953 TAATAAATGAATGACCAAGACTTTAATTAATAAATAAATAAATAAATAAATAAATAAATA 1012
Db 26073 TAATAAATGAATGATCAGACATTTAATTAATAAATAAATAAATAAATAAATAAATAA 26014
Qy 1013 AGTTATATCTGGATTTTCTCGGAAATAGGAACTACTTTTATATGATGATGTTT 1072
Db 26013 AAATTATGCTGGATTTTCTCGGAAATAGGAACTACTTTTATATGATGATGTTT 25954
Qy 1073 TTATGCAATTAGCATTTGATTTCTGTTCCAGGAATACTTTCTTAAGCAATATGTA 1132
Db 25953 TTATGCAATTAGCATTTGATTTCTGTTCCAGGAATACTTTCTTAAGCAATATGTA 25894
Qy 1133 GATATTAAAGATTAATAATCTAATGTATAAATAAATAAATAAATAAATAAATAAATA 1172
Db 25893 GATATTAAATATTAATAATCTAATGTATTGTAATAAATAAATAAATAAATAAATAA 25854

RESULT 5
CFRAB9
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

CFRAB9
Canine rab9 mRNA for ras-related GTP-binding protein.
X56386
X56386.1 GI:2189
GTP-binding protein; protein transport; ras gene; ras gene family.
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
1
Chavrier, P., Vingron, M., Sander, C., Simons, K. and Zerial, M.
Molecular cloning of YP1/SEC4-related cDNAs from an epithelial
cell line
Mol. Cell. Biol. 10 (12), 6578-6585 (1990)
PUBMED
2123294
REFERENCE
2 (bases 1 to 1001)
Lombardi, D., Soldati, T., Riederer, M.A., Goda, Y., Zerial, M. and
```

```
Pfeffer, S.R.
Rab9 functions in transport between late endosomes and the trans
Golgi network
EMBO J. 12 (2), 677-682 (1993)
93178443
8440258
3 (bases 1 to 1001)
Zerial, M.
Direct Submission
Submitted (17-AUG-1990) Zerial M., EMBL, Meyerhofstrasse 1, 6900
Heidelberg, FRG
FEATURES
source
1..1001
/organism="Canis familiaris"
/mol_type="mRNA"
/strain="cocker spaniel"
/db_xref="taxon:9615"
/clone="15"
/cell_line="MDCK"
1..1001
/gene="rab9"
<1..519
/gene="rab9"
<1..519
/gene="rab9"
/codon_start=1
/protein_id="CAA39797.1"
/db_xref="GI:2190"
/db_xref="SWISS-PROT:P24408"
/translations="NKFDTQLPHTIGVFLNKDLEVDGHPVTMQIWDVTAQGRFRSLR
TPYRSDCLLUTFSVDDSQNSUNWKEFIYADVKEPSPFPVILGNKIDISER
QVSTESAQAWCDNDGDPYFETSAKDNTVAAFEEAVRRLVATSDHLLIQDTVS
LHKPKPSSSCC"
<1..516
/gene="rab9"
/product="rab9 GTP-binding protein"
BASE COUNT 338 a 174 c 210 g 279 t
ORIGIN
Query Match 60.7%; Score 713.6; DB 4; Length 1001;
Best Local Similarity 87.1%; Pred. No. 1.4e-131;
Matches 881; Conservative 0; Mismatches 109; Indels 22; Gaps 8;
Qy 174 AATAAGTTTGATACCCAGCTCTTCCATCAATAGGTGTGGAATTTTAAATAAAGATTG 233
Db 1 AATAAGTTTGATACCCAGCTCTTCCATCAATAGGTGTGGAATTTTAAATAAAGATTG 60
Qy 234 GAAGTGGATGGACATTTTGTACCATGCAGATTGGGACACGCGAGGTCAAGAGCGATT 293
Db 61 GAGTGGATGGACATTTTGTACCATGCAGATTGGGACACAGCGGTCAAGAGCGATT 120
Qy 294 CGAAGCTGAGGACACCATTTTACAGAGTTCTGACTGTGCTGCTTACTTTAGTGT 353
Db 121 AGAAGCTGAGGACGCGCTTTTACAGAGTTCTGACTGTGCTGCTTACTTTAGTGT 180
Qy 354 GATGATTTCAGAGCTTCAGAACTTAAGTAAGTGAAGAAAGAAATTCATATATTATGCA 413
Db 181 GATGATTTCAGAGCTTCAGAACTTAAGTAAGTGAAGAAAGAAATTCATATATTATGCA 240
Qy 414 GATGTAAGAGAGCTGAGAGCTTTCTTTTGTGATTTCTGGGTAAACAAGATTGACATAAG 473
Db 241 GATGTAAGAGAGCGCGAAAGCTTTCTTTTGTGATTTTGGGCAACAAGATCGACATAAG 300
Qy 474 GAAGCGAGGTGTCTACAGAAAGCCCAAGCTTGGTGCAGGGAACAAGCGGACTATCCT 533
Db 301 GAAGCAAGTGTCTTACAGAAAGCCCAAGCTTGGTGCAGGGAACAAGCGGACTATCCT 360
Qy 534 TATTTTGAACAAAGTGAAGAAAGATGCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTT 593
Db 361 TACTTTGAACAAAGTGAAGAAAGATGCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTT 420
Qy 594 CGAAGATTTCTGCTACCGGATAGGTGATGATCATTTTGTTCACAGACAGACAGTCAAT 653
```

[illegible]

FEATURES
Source

Location/Qualifiers

```

1...130407
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="RP11-11"
/clone="RP11-809H9"
1..1423
/note="assembly_name:Contig10"
1524..3282
/note="assembly_name:Contig11"
3383..5435
/note="assembly_name:Contig12"
5536..7716
/note="assembly_name:Contig13"
7817..9788
/note="assembly_name:Contig14"
9889..11982
/note="assembly_name:Contig15"
12083..14414
/note="assembly_name:Contig16"
14515..16518
/note="assembly_name:Contig17"
16819..18797
/note="assembly_name:Contig18"
18898..20919
/note="assembly_name:Contig19"
21020..23820
/note="assembly_name:Contig20"
23921..26571
/note="assembly_name:Contig21"
26672..28476
/note="assembly_name:Contig22"
28577..31316
/note="assembly_name:Contig23"
31417..33788
/note="assembly_name:Contig24"
33889..36379
/note="assembly_name:Contig25"
36480..40039
/note="assembly_name:Contig26"
40140..45356
/note="assembly_name:Contig27"
45457..50509
/note="assembly_name:Contig28"
50610..55122
/note="assembly_name:Contig29"
55223..59638
/note="assembly_name:Contig30"
59739..64842
/note="assembly_name:Contig31"
64943..73926
/note="assembly_name:Contig32"
74027..86109
/note="assembly_name:Contig33"
86210..104005
/note="assembly_name:Contig34"
104106..127523
/note="assembly_name:Contig35"
127624..128816
/note="assembly_name:Contig6"
128917..130407
/note="assembly_name:Contig9"
BASE COUNT 39124 a 25068 c 24708 g 38792 t 2715 others
ORIGIN
Query Match 59.3%; Score 696.6; DB 2; Length 130407;
Best Local Similarity 82.4%; Pred. No. 1.2e-128;
Matches 930; Conservative 0; Mismatches 164; Indels 35; Gaps 10;
QY 59 AGGGTCTTGAAGCTTTGAGATTACAAATGGCAGGAAATCATCACTTTTAAAGTAAT 118
Db 38569 AGAGTTCTTGAGCTTTGAGATTAAATGTTGGCAGTAATTCATCACTTTTAAAGTAAT 38510

```


RESULT 7	
AC022486	
LOCUS	156357 bp DNA linear PRI 09-MAY-2001
DEFINITION	Homo sapiens BAC clone RP11-569J3 from Y, complete sequence.
ACCESSION	AC022486
VERSION	AC022486.4 GI:13435294
KEYWORDS	HTG
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Bukaryova; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 156357) Toward a complete human genome sequence Genome Res. 8 (11), 1097-1108 (1998) 99063792 9847074
AUTHORS	Scott, K., Maupin, R. and Swift, G.
TITLE	The sequence of Homo sapiens BAC clone RP11-569J3
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 156357) Waterston, R.H.
AUTHORS	Direct Submission
JOURNAL	Submitted (04-FEB-2000) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	4 (bases 1 to 156357) Waterston, R.H.
AUTHORS	Direct Submission
JOURNAL	Submitted (23-MAR-2001) Genome Sequencing Center, Washington University School of Medicine, 444 Forest Park Parkway, St. Louis, MO 63108, USA
REFERENCE	5 (bases 1 to 156357) Waterston, R.
AUTHORS	Direct Submission
JOURNAL	Submitted (09-MAY-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA On Mar 23, 2001 this sequence version replaced gi:7684584.
COMMENT	----- Center: Genome Center Center: Washington University Genome Sequencing Center Center code: WUGSC Web site: http://genome.wustl.edu/gsc Contact: sapiens@watson.wustl.edu ----- Summary Statistics ----- Center project name: H_NH0565J03 -----
NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.	
This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.	
MAPPING INFORMATION: The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.	
SOURCE INFORMATION: THE RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,	

REFERENCE

5 (bases 1 to 161879)

AUTHORS

Waterston,R.

TITLE

Direct Submission

JOURNAL

Submitted (09-MAY-2001) Department of Genetics, Washington

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

COMMENT

On Mar 23, 2001 this sequence version replaced gi:9211335.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>Contact: sapiens@wustl.wustl.edu

----- Summary Statistics

----- Center project name: H_NH0100J21

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skalecky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenoe, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-79J10; the clone sequenced to the right is RP11-245K4, 200 bp overlap. Actual start of this clone is at base position 1 of RP11-100J21; actual end is at base position 66964 of RP11-245K4.

FEATURES

source

1. .161879
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="y"
 /map="y"
 /clone="RP11-100J21"
 /clone_lib="RPCT-11"

repeat_region

1. .1185

/rpt_family="L1"

repeat_region

1183. .1513

/rpt_family="L1"

repeat_region

1679. .1942

/rpt_family="MER1_type"

repeat_region

2125. .3135

/rpt_family="ERV1"

repeat_region

3116. .3412

/rpt_family="ERV1"

repeat_region

3450. .3866

/rpt_family="ERV1"

repeat_region

3867. .6977

repeat_region
 6986. .7295
 /rpt_family="ERV1"
 repeat_region
 7297. .7625
 /rpt_family="ERV1"
 repeat_region
 9507. .9570
 /rpt_family="MER1_type"
 11143. .11817
 /rpt_family="L1"
 repeat_region
 12324. .12843
 /rpt_family="L1"
 repeat_region
 12874. .13922
 /rpt_family="ERV1"
 repeat_region
 12885. .13929
 /rpt_family="ERV1"
 repeat_region
 13051. .13264
 /rpt_family="L1"
 repeat_region
 13381. .13619
 /rpt_family="L1"
 repeat_region
 13828. .14306
 /rpt_family="L1"
 repeat_region
 15532. .15662
 /rpt_family="ERV1"
 repeat_region
 15740. .16044
 /rpt_family="ERV1"
 repeat_region
 16045. .16136
 /rpt_family="ERV1"
 repeat_region
 16178. .16239
 /rpt_family="ERV1"
 repeat_region
 16728. .17201
 /rpt_family="ERV1"
 repeat_region
 17188. .17253
 /rpt_family="ERV1"
 repeat_region
 17243. .17602
 /rpt_family="ERV1"
 repeat_region
 17746. .18064
 /rpt_family="L1"
 repeat_region
 18085. .18235
 /rpt_family="L1"
 repeat_region
 18303. .18876
 /rpt_family="ERVK"
 repeat_region
 18927. .19413
 /rpt_family="ERVK"
 repeat_region
 19419. .19628
 /rpt_family="L1"
 repeat_region
 21508. .21759
 /rpt_family="Alu"
 repeat_region
 22725. .22868
 /rpt_family="ERV1"
 repeat_region
 22868. .23335
 /rpt_family="ERV1"
 repeat_region
 23350. .24076
 /rpt_family="ERV1"
 repeat_region
 24273. .24690
 /rpt_family="ERV1"
 repeat_region
 24702. .24823
 /rpt_family="ERV1"
 repeat_region
 24886. .25175
 /rpt_family="Alu"
 repeat_region
 25186. .27543
 /rpt_family="ERV1"
 repeat_region
 27654. .27697
 /rpt_family="ERV1"
 repeat_region
 28186. .28421
 /rpt_family="ERV1"
 repeat_region
 28440. .28804
 /rpt_family="ERV1"
 repeat_region
 28773. .29267
 /rpt_family="ERV1"
 repeat_region
 29270. .31939
 /rpt_family="ERV1"
 repeat_region
 31933. .32211
 /rpt_family="ERV1"

[illegible]

TITLE
JOURNAL
Submitted (25-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS
7 (Bases 1 to 166436)
Waterston,R.
TITLE
JOURNAL
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
On Apr 23, 2000 this sequence version replaced gi:6524293.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@wustl.edu
----- Summary Statistics

Center project name: H_NH0506M09

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skaletsky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-221K4. Actual start of this clone is at base position 1 of RP11-506M9; actual end is at base position 166436 of RP11-506M9.

FEATURES

source

1. .166436
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="Y"
/map="Y"
/clone="RP11-506M9"
/clone_lib="RP11-11"
195. .923
/rpt_family="L1"
924. .1077
/rpt_family="L1"
1060. .1830
/rpt_family="L1"
1879. .1918
/rpt_family="L1"
1905. .3257
/rpt_family="L1"

repeat_region
3260. .3393
/rpt_family="Alu"
repeat_region
3618. .3694
/rpt_family="MIR"
repeat_region
4250. .4511
/rpt_family="Alu"
repeat_region
4512. .5804
/rpt_family="L1"
repeat_region
5804. .7682
/rpt_family="L1"
repeat_region
8188. .8294
/rpt_family="L2"
repeat_region
9429. .9499
/rpt_family="L2"
repeat_region
9826. .9872
/rpt_family="(A)n"
repeat_region
9962. .10035
/rpt_family="L2"
repeat_region
10596. .10684
/rpt_family="GA-rich"
repeat_region
12836. .13209
/rpt_family="MaLR"
repeat_region
14667. .14689
/rpt_family="AT-rich"
repeat_region
16692. .16712
/rpt_family="AT-rich"
repeat_region
17516. .17584
/rpt_family="MERS3"
repeat_region
17590. .17667
/rpt_family="MERS3"
repeat_region
17814. .18112
/rpt_family="Alu"
repeat_region
19421. .19650
/rpt_family="L1"
repeat_region
19889. .19813
/rpt_family="Retroviral"
repeat_region
19815. .20107
/rpt_family="Retroviral"
repeat_region
20410. .20861
/rpt_family="Retroviral"
repeat_region
21083. .25178
/rpt_family="Retroviral"
repeat_region
25196. .27960
/rpt_family="Retroviral"
repeat_region
28006. .28520
/rpt_family="Retroviral"
repeat_region
28328. .28905
/rpt_family="L2"
repeat_region
29157. .29395
/rpt_family="Alu"
repeat_region
29396. .29438
/rpt_family="(TAA)n"
repeat_region
30965. .30996
/rpt_family="AT-rich"
repeat_region
31776. .32149
/rpt_family="MaLR"
repeat_region
33205. .33489
/rpt_family="Alu"
repeat_region
35802. .35832
/rpt_family="AT-rich"
repeat_region
36311. .36389
/rpt_family="(TTAAA)n"
repeat_region
36757. .36827
/rpt_family="MERS3"
repeat_region
38147. .38551
/rpt_family="Retroviral"
repeat_region
38552. .38576
/rpt_family="(TC)n"
repeat_region
38577. .38633
/rpt_family="Retroviral"
repeat_region
38634. .39399
/rpt_family="Retroviral"
repeat_region
39400. .39459

```
/rpt_family="MER1_type"
39460. .40129
/rpt_family="Retroviral"
40121. .41902
/rpt_family="Retroviral"
41331. .42280
/rpt_family="Retroviral"
42281. .42759
/rpt_family="Retroviral"
42961. .43305
/rpt_family="L1"
43410. .43703
/rpt_family="Alu"
44044. .44848
/rpt_family="L1"
```

Query Match 58.1%; Score 682.8; DB 9; Length 166436;
Best Local Similarity 81.4%; Pred. No. 6.4e-126;
Matches 904; Conservative 0; Mismatches 162; Indels 44; Gaps 8;

```
QY 59 AGGGTCTTGAAGCTTTTGAGATTAAATGCGAGGAAATCATCTTTTAAAGTAAT 118
Db 148862 AGGGTCTTGAAGCTTTGCAATTAACAGTGGCAGGAATTCATCTTTTAAAGTAAT 148921

QY 119 TCTCCTTGGAGATGGTGGAGTGGGAGAGTTCACCTATGAACAGATATGTAACATAA 178
Db 148922 TCTCCTTGGAGATGGTGGGAGTGGGAGAGTTCACCTATGAACAGATATGTAACATAA 148981

QY 179 GTTTGATACCCAGCTCTTCATACAAATAGGTGTGGAATTTTAAATAAGATTCGAAGT 238
Db 148982 GTTTGCACTCAGCTGTTCATGCAATAGATGTGGAATTTTAAATAA--GATTTGAAG 149039

QY 239 GATGACATTTTGTACCATGACAGATTTGGGACACGGCAGGTGAGGACGATTCGGAAG 298
Db 149040 GGTGGACATTTTGTACCATGACAGATTTGGGACATGATTTGTGACATGGCAGGTTCGGAAG 149099

QY 299 CTTGAGGACACCATTTTACAGAGTCTGACTGCTGCTGCTTACTTTTGTAGTGTGATGA 358
Db 149100 CTTGAGGACACCATTTTACAGAGTCTGACTGCTGCTGCTTACTTTTGTAGTGTGATGA 149159

QY 359 TTCACAAAGCTCCAGAACTTAAGTAACCTGGAAGAAAGAAATTCATATATATGTCAGATGT 418
Db 149160 TTTAGGAAGCTCCAGAACTTAAGTAATTTG----- 149189

QY 419 GAAAGAGCTGAGAGCTTCTCTTTGTGATTTCTGGTAAACAGATTCATTAAGGACGAG 478
Db 149190 GAAAGAACTGAGCGATTTCTTTGTAGTTCTGCAATTCAGGACTGGCATTAACGGAAG 149249

QY 479 GCAGTGTCTACAGAAAGCCCAAGCTTGTGTCAGGGAACACGGCGACTATCCTTTATTT 538
Db 149250 GCAGGTGTTTATAGAAGAA--CAGCTGCAATTATGCAGGAGGACATGACTACCTTACTT 149308

QY 539 TGAACAAGTGCMAAGATGCCACAAATGTGGCAGCAGCTTTGAGGAAGCGGTTTCGAAG 598
Db 149309 TGCAACAGTGCMAAGATGCCACAAATGTGGCAGCAGCTTTGAGGAAGCGGTTTCGAAG 149368

QY 599 AGTTCTTGTCCAGGAGATAGTTCAGATCATTGATTCAGACAGACACAGTCAATCTTCA 658
Db 149369 AATTTCTTGTCACTGTGATATGTCAGATCCTTGATGACACAGGCAATGTCACCTTCA 149428

QY 659 CCGAAGCCCAAGCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
Db 149429 CCAAGAGCCCAAGCTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 149488

QY 715 TAACCAACTCACACATATACAAATCAACATGCGGATGGGAGAGAGATTTAGGTTTG 774
Db 149489 TAACCAACTCACACATATACAAATCAACATGCGGATGGGAGAGAGATTTAGGTTTG 149548

QY 775 CAGCAGTGTATCATCTTACTATAATAAATAAATAATGT-----TGCTGCTTCTATTAGTTGG 830
Db 149549 TAGCAGTGTATTTACTTACTATAATAAATAAATAATGT-----TGCTGCTTCTATTAGTTGG 149608

QY 831 TGGGAGAGGGACATCACTCTCTGGAGGAATATATTTACTCAATATGCGACCTTACA 890
```

```
Db 149609 TGGGAGATGGACATATCCACTCATGGAGGATATATTTTCTCAATATGAACCTTACA 149668
QY 891 TTTATAAATTTGTAACAGTGTGCTAATAACGTTTCTTTTAATTTAAATATGTAAGTTGCAGA 950
Db 149669 TTTATAAATTTGTAATGTTGTTCTAA--AAAGTCTCTTTGATTTACACATGCAATTAGAGA 149727
QY 951 GCTAATAATGAATGACCAAGACTTTAATATTAATAAATAAATAAATAAATAAATAAATAAATA 1010
Db 149728 GCTAATAATGAGAGGACCAAGACTTTAATATTAATAAATAAATAAATAAATAAATAAATAAATA 149787
QY 1011 AGAAGTTATCTTG--GATTTTCTCGGAAATGGAAGACTTCTTTTATATATGTAATG 1069
Db 149788 AGAATTTATCTTGAAATTTTCTCGGAAATGGAAGACTTCTTTTATATGTAATG 149847
QY 1070 -TTTTATGCAATTAGCAATGTAATTTCTGTCAGGAAATACTTTCTTAAGCAATAAT 1128
Db 149848 ATTTTGTGCAATTAACATGTAATTTCTGTCAGGAAATGATTTCTTAAGCAATAAT 149907
QY 1129 GTTAGATATTAAAGATTAAATCTAAATGTA 1158
Db 149908 GTTGGATATGAAGATTAGATCTTAACATA 149937

RESULT 11
AC007965/c
LOCUS
DEFINITION Homo sapiens BAC clone RP11-245K4 from Y, complete sequence.
ACCESSION AC007965
VERSION AC007965.3 GI:10047995
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
9847074
REFERENCE
2 (bases 1 to 182083)
Du,H., Stoneking,T. and Gregory,S.
The sequence of Homo sapiens BAC clone RP11-245K4
Unpublished
3 (bases 1 to 182083)
Waterston,R.H.
Direct Submission
Submitted (03-JUL-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 182083)
Waterston,R.H.
Direct Submission
Submitted (10-SEP-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 182083)
Waterston,R.
Direct Submission
Submitted (30-SEP-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 10, 2000 this sequence version replaced gi:5523821.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0245K04
```

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The position of this clone was established as part of a collaboration between the Human Chromosome Y Mapping Project (Tomoko Kawaguchi, Helen Skalecky, Laura G. Brown, Steve Rozen, and David C. Page at the Whitehead Institute for Biomedical Research, Cambridge MA) and the Washington University Genome Sequencing Center, St. Louis MO.

SOURCE INFORMATION:

The RPCT-11 human BAC library was made from the blood of one male donor, as described by Ooegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanesi, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-100J21; the clone sequenced to the right is RP11-270H4, 200 bp overlap. Actual start of this clone is at base position 1 of RP11-245K4; actual end is at base position 15260 of RP11-270H4.

The sequence RP11-245K4 from base positions 105307 to 105967 and 141470 to 141790 is represented by sequence derived from PCR of clone dna.

FEATURES

source

```
1. .182083
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="Y"
   /map="Y"
   /clone="RP11-245K4"
   /clone_lib="RPCT-11"
```

repeat_region

```
1. .65
   /rpt_family="L1"
```

repeat_region

```
3688. .3775
   /rpt_family="L1"
```

misc_feature

```
9345. .9559
```

misc_feature

```
9372. .9555
   /note="similar to EST AI654951 (NID:G4738930) wb65g09.x1"
```

misc_feature

```
10775. .10876
   /note="similar to EST AI767654 (NID:G5234163) wh38b10.x1"
```

misc_feature

```
11063. .11277
   /note="similar to EST AI767654 (NID:G5234163) wh38b10.x1"
```

repeat_region

```
11875. .12075
   /note="similar to EST AI767654 (NID:G5234163) wh38b10.x1"
```

misc_feature

```
12065. .12386
   /rpt_family="MER1_type"
```

repeat_region

```
12856. .13310
   /note="similar to EST B5093625 (NID:G6484077) "
```

repeat_region

```
13650. .13896
   /rpt_family="MaLR"
```

repeat_region

```
13919. .14086
   /rpt_family="Alu"
```

repeat_region

```
14857. .14991
   /rpt_family="MER1_type"
```

repeat_region

```
15334. .15397
   /rpt_family="MER1_type"
```

```

/rpt_family="L2"
15406. .15530
/rpt_family="MIR"
15762. .15841
/rpt_family="Alu"
15840. .15972
/rpt_family="Alu"
16714. .16830
/note="similar to EST AA020179 (NID:G1483961) mh51g02.r1"
16716. .16830
/note="similar to EST AA874484 (NID:G2979173) vx80h07.r1"
16716. .16829
/note="similar to EST AI918053 (NID:G5637908) w138g01.x1"
16716. .16829
/note="similar to EST AW207673 (NID:G6507169) "
16716. .16829
/note="similar to EST AW770235 (NID:G7702298) h179b07.x1"
16716. .16829
/note="similar to EST T89537 (NID:G718050) ye04c07.r1"
16717. .16830
/note="similar to EST AI928109 (NID:G5664073) wp11b12.x1"
16717. .16830
/note="similar to EST AW140982 (NID:G6160750) "
16717. .16829
/note="similar to EST AW027866 (NID:G5886622) ws62b12.x1"
16722. .16830
/note="similar to EST AA799342 (NID:G4131386) "
16736. .16830
/note="similar to EST AI819740 (NID:G5438819) wj11g03.x1"
16736. .16829
/note="similar to EST AI312610 (NID:G4018215) qp77b06.x1"
16751. .16829
/note="similar to EST N40075 (NID:G1163620) yx98e03.r1"
18145. .18569
/rpt_family="MaLR"
18606. .19251
/note="similar to EST AA203531 (NID:G1799242) zx56g09.r1"
18782. .18932
/note="similar to EST AA799342 (NID:G4131386) "
18782. .18932
/note="similar to EST AA854621 (NID:G2942159) aj76f06.s1"
18782. .18932
/note="similar to EST AA874484 (NID:G2979173) vx80h07.r1"
18782. .18932
/note="similar to EST AI819740 (NID:G5438819) wj11g03.x1"
18782. .18932
/note="similar to EST AI928109 (NID:G5664073) wp11b12.x1"
18782. .18932
/note="similar to EST AW140982 (NID:G6160750) "
18782. .18929
/note="similar to EST AA020179 (NID:G1483961) mh51g02.r1"
18784. .18932
/note="similar to EST AA715483 (NID:G2727757) nv53f10.r1"
18784. .18932
/note="similar to EST AI312610 (NID:G4018215) qp77b06.x1"
18784. .18932
/note="similar to EST AI918053 (NID:G5637908) w138g01.x1"
18784. .18932
/note="similar to EST AW027866 (NID:G5886622) ws62b12.x1"
18784. .18932
/note="similar to EST AW207673 (NID:G6507169) "
18784. .18932
/note="similar to EST AW770235 (NID:G7702298) h179b07.x1"
18784. .18932
/note="similar to EST N40075 (NID:G1163620) yx98e03.r1"
18784. .18932
/note="similar to EST T89537 (NID:G718050) ye04c07.r1"
18798. .18932
/note="similar to EST AI033300 (NID:G3254253) ow74e10.s1"
19378. .19473
/note="similar to EST AA799342 (NID:G4131386) "
19378. .19473
/note="similar to EST T89537 (NID:G718050) ye04c07.r1"

```



```
misc_feature 19378..19463
/notes="similar to EST AA715483 (NID:92727757) hv53f10.r1"
misc_feature 19378..19463
/notes="similar to EST AA854621 (NID:92942159) aj76f06.s1"
misc_feature 19378..19463
/notes="similar to EST AI312610 (NID:94018215) qp77b06.x1"
misc_feature 19378..19463

Query Match
Best Local Similarity 81.4%; Pred. No. 6.2e-126; Length 182083;
Matches 904; Conservative 0; Mismatches 162; Indels 44; Gaps 8;

Qy 59 AGGTTCTTGAAGTTTGGAGTAAACAATGGCAGGAAATCATCACTTTTAAAGTAAT 118
Db 31142 AGGGTTCTTGAAGTTTGGAGTAAACAATGGCAGGAAATCATCACTTTTAAAGTAAT 31083
Qy 119 TCTCCTTGGAGATGGTGGAGTTCACCTATCAACAGATATGTAACTAATAA 178
Db 31082 TCTCTTGGAGATGGTGGAGTTCACCTATCAACAGATATGTAACTAATAA 31023
Qy 179 GTTTGATACCCAGCTCTTCATCAATAGGTGTGGAATTTTAAATAAAGATTTGGAAGT 238
Db 31022 GTTTGGCACTCAGCTGTTCATCAATAGATGTGGAATTTTAAATAA--GATTTGAAG 30965
Qy 239 GGATGGACATTTTGTACCATGAGATTTGGGACAGCGCAGGTCCGAGGATTCGGAAG 298
Db 30964 GGTTGGACATTTTGTACCATGAGATTTGTGACATGGCAGGTCCGAGGATTTCTGAAG 30905
Qy 299 CTTGAGGACACCATTTTACAGAGTTTCTGACTCTGCTGCTTACCTTTTAGTGTGATCA 358
Db 30904 CTTGAGGACACCATTTTACAGAGTTTCTGACGTTCTGCTTATTTTAGAATCAATGA 30845
Qy 359 TCCAAAGCTTCCAGACTTAAGTAACCTGGAAGAAAGAAATTCATATATATGCGAGATGT 418
Db 30844 TTTAGGAAGCTTCCAGAACTTAAGTAATGT----- 30815
Qy 419 GAAAGAGCTTGAGAGCTTTCCTTTTCTGATCTGGGTAAACAATGATCAATGCGAAGC 478
Db 30814 GAAAGAGCTTGAGAGCTTTCCTTTTCTGATCTGGGTAAACAATGATCAATGCGAAGC 30755
Qy 479 GCAGGTGTCTACAGAAAGACCCCAAGCTTGGTCAGGAGCAAGCGGACTATCCTTATTT 538
Db 30754 GCAGGTGTCTTATAGAGAA--CAGCTGCATTTATGACAGAGGACAGTGACTACCTTACTT 30696
Qy 539 TGAACAAGTGCAAAAGATGCCAATATGTGGCAGCAGCTTTGAGGAAGCGGTTGCGAAG 598
Db 30695 TGAACAAGTGCAAAAGATGCCAATATGTGGCAGCAGCTTTGAGGATGGGGTTCAAAA 30636
Qy 599 AGTTCTTGCTACCGAGATAGTGCAGATCATTTGATTCAGACAGACACAGTCAATCTTCA 658
Db 30635 AATCTTGTGACTGTGGATATGTCAGATCATTTGATGAACACAGGCGATGGTCAACCTTCA 30576
Qy 659 CCGAAGCCCAAGCTAGCTCATCTTCTGTTGATTTAG----ATTGTTGATGCATTC 714
Db 30575 CCAAAAGCCCAAGCTGGCAGATCTTGTATTGATTTAGATATAGGATGAAGCAGATGCATGC 30516
Qy 715 TACCAACTCACACATATACAGAAATCAACATGGGGATGGAGAGAGAAATAGGCTTGG 774
Db 30515 TAACCAATTCACACATATACAGAAATCAACATGGGATGGAGAGAGAAATAGGCTTGG 30456
Qy 775 CAGCAGTGTATCATCTACTAATAAATAAATAAATAATGT----TGCTGCTTCATTTAGTTGG 830
Db 30455 TACAGTGTATTTACTTACTAATAAAGATTCACAAATGTATAGCCCTGTGTCATTTAGTTGG 30396
Qy 831 TGGAGAGGAGACATCCACTCTTGGAGAGATATATTTACTCAATATGCGACCTTACA 890
Db 30395 TGGAGATGGGACATATCCACTCATGGAGAAATATATTTTCTCAATAATGAACCTTACA 30336
Qy 891 TTTATAAATGTAAAGTGTCTTAATAAGCTTTCTTTAATTTTAAATATATGATGTCAGA 950
Db 30335 TTTATAAATGTAAAGTGTCTTAATAAGCTTTCTTTAATTTTAAATATATGATGTCAGA 30277
Qy 951 GCTAATAAATGAATGACCAAGCACTTTAATTAATAAATAAATAAAGAACTTGACTATTTCT 1010
```

```
Db 30276 GCTAATAAATGAGAGGACCAAGACTTTAATTAATTAATTAAGATAAGATACCTGACTATCT 30217
Qy 1011 AGAGTTTATCTTG-GATTTTCTCGGAAAATGCGAAGACTACTTTTATATGTTGATG 1069
Db 30216 AGAATTTATCTTGAATTTTCTCGGAAAATGCGAAGACTACTTTTCTATGTACATG 30157
Qy 1070 -TTTTATGCAATTAGCATTTGTTTCTGTTTCAGGAAAATACTTTCTTAAAGCAATAAT 1128
Db 30156 ATTTTGTGCAATTAACATTTGTTTCTGTTTCAGGAAAATGATTTTCTTAAAGCAATAAT 30097
Qy 1129 GTTAGATTAATAAGATTAAATCTAATGTA 1158
Db 30096 GTTGTATGAAGATTAGAATCTTAACATA 30067

BC008160 1251 bp mRNA linear ROD 16-APR-2003
Mus musculus RAB9, member RAS oncogene family, mRNA (CDNA clone
MGC:11411 IMAGE:3964134), complete cds.
ACCESSION BC008160
VERSION BC008160.1 GI:14919427
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1251)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,K., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Sheezy,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raja,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Buckfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1251)
Strausberg,R.
Direct Submission
Submitted (22-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Naravati,
A.N., Gibbs, R.A.
```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 17 Row: K Column: 20
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 9790226.

FEATURES

source

Location/Qualifiers
 1..1251
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVE/N"
 /db_xref="taxon:10090"
 /clone="MGC:11411 IMAGE:3964134"
 /tissue type="mammary tumor. Metallothionien-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
 /clone.lib="NCI CGAP_Mam1"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..1251
 /gene="Rab9"
 /note="synonyms: Sid6061p, 2410064E05Rik"
 /db_xref="LocusID:56382"
 /db_xref="MGI:1890695"
 186..731

gene

CDs
 /codon_start=1
 /product="RAB9, member RAS oncogene family"
 /protein_id="RAH08160.1"
 /db_xref="GI:14919428"
 /db_xref="LocusID:56382"
 /translation="MAGKSSLPKILLDGVGVKSSLMNRVVTNPKDSOLFHTIGVEF LNKOLEVDGVHPTVQIWDYTAGQERPSRLRFFYRGSBCLLTFSDVDSQFQNLNWK KEFIYADKVPFSPFVILGNKTDIKERQVSTEEAQAQWCKNDGDFYFETSANDSTN VAAAFEEAVRILATEDSEHLIQTDTVLRHKPKNSOC"
 390 a 247 c 270 g 344 t

Query Match 56.6%; Score 664.6; DB 10; Length 1251;

Best Local Similarity 78.5%; Pred. No. 7.1e-122; Mismatches 214; Indels 30; Gaps 7;

Matches 893; Conservative 0;

36 TGCACTCTCTCAACACCCCTAGGAGGGTCTTGAAGCTTTTCAGATTAAACATGGCAGGA 95

135 TGCTTGAAGTGAACCTAAGGATTCGAATCAAACTTACATTATTAATGACGAGGA 194

96 AAATCATCATCTTTTAAAGTAATTCCTCTTGAGAGATGTGGAGTTGGAGAGAGTTCACTT 155

195 AAATCGTCTCTTTTAAATTAATTCCTCTTGAGAGATGTGGAGTTGGAGAGAGTTCTCTT 254

156 ATGAACAGATATGTAACTAATAAGTTTGATACCCAGCTCTTCCATACAAATAGGTGTGAA 215

255 ATGAACAGATATGTAACTAATAAGTTTGATACCCAGCTCTTCCATACAAATAGGTGTGAA 314

216 TTTTAAATAAAGATTTGGAATGGATGGACATTTTGTATACCATGACAGATTGGGACAG 275

315 TTTTAAATAAAGATCTGGAGGTGGACGACATTTTGTATACCATGACAGATTGGGACACA 374

276 GCAGTCCAGGCGGATTCGAGCTGAGGACACCATTTTACAGAGTTCTGACTGTGCG 335

375 GTGGCCAGAGACGGTTCGAGCTGAGGACGACATTTTACCAGGTTCTGACTGTGCG 434

336 CTGCTTACTTTTGTGATGATTCACAAAGCTTCCAGAACTTAAAGTAACCTGGAAGAAA 395

435 CTGCTTACTTTTGTGATGATTCACAGAGCTTCCAGAACTTAAAGTAACCTGGAAGAAA 494

396 GAATTCATATATATGACAGATGTGAAGAGCTTGAGAGCTTCTTTTGTGATTCGGGT 455

495 GAGTTTCATATATATGACAGATGTGAAGAGCTTGAGAGCTTCTTTTGTGATTTGGGC 554

456 AACAGATTGACATTAAGCAACGGCAGGTGTCTACAGAGAGAGCCCAAGCTTGGTGCAGG 515

555 AACAGACTGACATTAAGCAACGGCAGGTGTCTACAGAGAGAGCCCAAGCTTGGTGCAGG 614

516 GACAAAGGCGCATATCTCTTTATTTTGAACAAGTGCAGAAAGATGCCACAAATGTGGCAGCA 575

615 GACAAAGGCGCATATCTCTTACCTTTGAAACAAGTGCAGAAAGATTCACCAATGTGCTGCT 674
 576 GCCTTTGAGAGAGCGGTTCGAAGAGTTCTTGTACCGAGGATAGTCAGATCATTTGATT 635
 675 GCCTTTGAGAGAGCGGTTCGAAGAGTTCTTGTACCGAGGATAGTCAGATCATTTGATT 734
 636 CAGACAGACACAGTCAATCTTCCACCAAGCCCAAGCCCTAGCTCATCTTGTGTTGATTG 695
 735 CAGACAGACACAGTCAATCTTCCACCAAGCCCAAGCCCTAGCTCATCTTGTGTTGATTG 794
 696 -----TTAGATTGTTGATGCAATTTTAAACCAACTCACATATACACAAAATCAACATGG 749
 795 CATTAAGGAGATAGTTGGTGCATTTTAAACCAACACACACACACACAC-----ACAGAA 847
 750 GGATGGAGAGAGATTAGCGTTTTCAGCAGGTGATCTACTACTAATA-AAATTAACCTA 808
 848 AGGTGGAGAAACAAATTTAGTATGACAGAGAGTTCAATTTACTAATAAATTCAGTTAAGC 907
 809 ATGTTGCTCTCTTCAATTTAGTTGGGAGAGGAGACATCCACTCTTGGAGGAATATATT 868
 908 ATATTGTTGCTCTCAATTTAGTGGTGGAGAGGACACTCCTC-----TGAGGAACTCTATT 964
 869 TACTCAATATGCGACCTTACATTTATTAATTTGTAACAGTTGCTTAATAACGTTTCTTTA 928
 965 TACTCAGTAATGCGACCTTCCACTTATAAATTTGTAATTTGTTCTTAAT---GTTTCTTTA 1021
 929 ATTTAAATATGTAAGTTGCGAGAGCTTAATAATGAATGCAAGAGCTTTAATAATAATAA 988
 1022 AATTAACAATTT-----AGTGTCTAATAAGATGACCAAGAGTACTGTTAATTC 1074
 989 AATAAGAAACTTGACTATCTAGAGT---TATCTGAGATTTTCTTCTGGGAAAATGG 1045
 1075 AGACGACAACTTGGCTATTCTAGAGTACACTTAGATTGTTTGGACCCACAGGAAATG 1134
 1046 AGAAGTACTTTTATATGTTGTTTATCAATAGCTAGCTTATCTGTTTCAGGG 1105
 1135 GAAATATCTTTTATATGTTGTTTATTTTATGTAATTTAGTATTTCTTCTGTTAGGA 1194
 1106 AAATACTTTCTTAAAGCAATAATGTTAGATATTAAAGATTAATAATCTAATGTAAAAA 1162
 1195 AAGTAAATTTCTACAGCAGTACTATTAAAGATTAATAATCTAATAAATAAATAAATAA 1251

RESULT 13

AB027290

LOCUS

DEFINITION

cds.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..1276

/organism="Mus musculus"

/mol_type="mRNA"

AB027290

Mus musculus

SID 99

mRNA

linear

ROD 29-SEP-1999

Mus musculus

small GTP binding protein, complete

cds.

AB027290

GI:5931615

small GTP binding protein.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Seki.N., Hattori.A., Hayashi.A., Kozuma.S., Muramatsu.M. and

Saito,T.

Mouse small GTP binding protein

Published Only in DataBase (1999)

2 (bases 1 to 1276)

Seki.N., Hattori.A., Hayashi.A., Kozuma.S., Muramatsu.M. and

Saito,T.

Direct Submission

Submitted (14-MAY-1999)

Toshiyuki Saito, National Institute of

Radiological Sciences, Genome Research Group; Inage-ku Anagawa

4-9-1, Chiba, Chiba 263-8555, Japan [E-mail:t_saito@nirs.go.jp,

Tel:81-43-201-3135, Fax:81-43-251-9818]

Location/Qualifiers

1..1276

/organism="Mus musculus"

/mol_type="mRNA"

```

/db_xref="taxon:10090"
1. .1276
/gene="SID 99"
216. .821
/gene="SID 99"
/codon_start=1
/product="small GTP binding protein"
/protein_id="BAA84709.1"
/db_xref="GI:5931616"
/translation="MACKSLFKILLGDSGVKSLMRYVTNKDSQLFHWIVGF
LNKLEVDVGHFVMOJWDTAGQERFSLRFPFVGRGDCCLLTFSDVDSQFQLSNWK
KEFYIADVKEPSFPVILGNKIDIKEROVSTEBQAWCKNDGDPYFETSAKDSIN
VAAFEAVARILLETDRSEHLIQTDVNLHRKPKNSCC"
BASE COUNT 384 a 264 c 274 g 354 t
ORIGIN

Query Match 56.2%; Score 660.2; DB 10; Length 1276;
Best Local Similarity 78.9%; Pred. No. 5.3e-121; Indels 37; Gaps 8;
Matches 899; Conservative 0; Mismatches 203;

QY 36 TGCATCTTCTCAACAACCCCTAGGAGGTTCTTGAAGCTTTTGAGATTAAACAATGGCAGGA 95
DB 165 TGCCTTGAAGTGAACCTACTAGGATTCGAATCAAACTTACATTATTAAATGGCAGGA 224
QY 96 AAATCATCATCTTTTAAAGTAATCTCTTGGAGATGGTGGAGTGGGAAGAGTTCACCT 155
DB 225 AATCGTCTCTTTTAAATAAATCTCTTGGAGATGGTGGAGTGGGAAGAGTTCCTT 284
QY 156 ATGAACAGATGATGAACTAATAGTTTGATACCCAGCTCTTCCATACAATAGTGTGGAA 215
DB 285 ATGAACAGATGATGAACTAATAGTTTGATACCCAGCTCTTCCACACATAGTGTGGAA 344
QY 216 TTTTAAATAAAGTTTGGAGTGGATGGACATTTGTTACCATGATGGAGTTGGGACAG 275
DB 345 TTTTAAATAAAGTCTGGAGTGGAGCGACATTTGTTACCATGATGGAGTTGGGACACA 404
QY 276 GCAGGTGAGGAGCCATCCGAAGCTCGAGACACCATTTTACAGAGTCTTCTGACTGCTG 335
DB 405 GCTGGCCAGACGCTCCGAGCTGAGAGCGCATTTTACCGGGTCTGACTGTGCT 464
QY 336 CTGCTTACTTTTGTAGTCGATGATCAAAAGCTTCAGAACTTAAAGTAACTGGAAGAAA 395
DB 465 CTGCTTACATTTAGTGTGATGATTCACAGAGCTTCAGAACTTAAAGTAACTGGAAGAAA 524
QY 396 GAATTCATATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 455
DB 525 GAGTTTATATATACGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 584
QY 456 AACAAAGATTGACATAAGCGAGCGAGTGTCTACAGAAAGCCCAAGCTTGGTGCAGG 515
DB 585 AACAAAGATTGACATAAGCGAGCGAGTGTCTACAGAAAGCCCAAGCTTGGTGCAGG 644
QY 516 GACACGGCGACTATCTTTATTTTAAAGTGGCAAGATGCGCAAAATGCGCAGCA 575
DB 645 GACACGGCGACTATCTTACTTTGAACAAAGTGGCAAGATGCGCAAAATGCGTGTCT 704
QY 576 GCCTTTGAGGAAGCGGTTCAAGAGTCTTGTACCGAGATAGGTCAGATCATTTGATT 635
DB 705 GCCTTTGAGGAAGCGGTTCAAGAGTCTTGTACCGAGATAGGTCAGATCATTTGATT 764
QY 636 CAGACGACACAGTCAATCTTCCGAAAGCCCAAGCTTCACTTCTGCTGTGATTG 695
DB 765 CAGACGACACAGTCAATCTTCCGAAAGCCCAAGCTTCACTTCTGCTGTGATTG 824
QY 696 -----TTAGATTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 749
DB 825 CATTAAAGGATAGTGTGATGATGATGATGATGATGATGATGATGATGATGATG 877
QY 750 GGATGAGAGAGAAATAGCGTTTGACGAGTGTATCATCTACTATAA-AAATTAAGTA 808
DB 878 AGGTGAGAGAAACAAATAGTATGACAGAGTTCATTTTACTAATAAATTCAGTTAACG 937
QY 809 ATGTTCTGCTTCAATAGTTGTTGGAGAGGACACATCCACTCTTGTGAGGAATATATT 868

```

```

DB 938 ATATTTGTCCTCATTTAGTCGTGGAGAGGACACTCCTC---TGAGGAATCTATT 994
QY 869 TACTCAATAATGGCACCCTTACATTTTATAAATTTGTAACAGTTGTCTTAATAACCTTTCTTTA 928
DB 995 TACTCAGTAATGGCAGCTTCCACTTATAAATTTGTAATTTGTTGCTAAAT---GTTTCTTTA 1051
QY 929 ATTAAATATGTAAGTTGCGAGAGCTAATAAATGAATGACCAAGACTTAAATTAATAA 988
DB 1052 AATTAATAACATT-----AGTGCTAATAAGATGCCAAGAACTGACTTTTACTGTAAATTC 1104
QY 989 AAATAAGAACTTGACTATTCTTAGAAGT---TATACCTTGGATTTTCTCTGGGAAATGG 1045
DB 1105 AGACGACAACTTGGCTATTCTTAGAAGTACACTTAGATTGTTTGCACCCACCAAGAAATG 1164
QY 1046 AGAACTACTTTTATATATGTTATGTTTATGCAATAGCATTTGTTCTTGGTTCAGGG 1105
DB 1165 GAAATTAATCTTTTATATGTTATATTTTATGTAATTTAGCATTTGTTCTTGGTCTAGGA 1224
QY 1106 AAATACTTTCTTAAAGCAATATGTTAGATATTAAGATTAATAATCTAATGTAAATAAA 1164
DB 1225 AAGTAATTTCTTACAGCAGTA-----CTAATAAGATTAATAATCTAATGCAAAAAA 1276

RESULT 14
AL6721174/c
LOCUS
DEFINITION
  Mouse DNA sequence from clone RP23-408K4 on chromosome X, complete
  sequence.
ACCESSION
  AL6721174
VERSION
  AL6721174.8 GI:24940538
KEYWORDS
  HTG.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 201642)
  Howden,P.
  Direct Submission
  Submitted (09-NOV-2002) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Nov 12, 2002 this sequence version replaced gi:21727414.
  ----- Genome Center
  Center: Wellcome Trust Sanger Institute
  Center code: SC
  Web site: http://www.sanger.ac.uk
  Contact: humquery@sanger.ac.uk
  -----
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  The following abbreviations are used to associate primary accession
  numbers given in the feature table with their source databases:
  Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
  on the WORMPEP database can be found at
  http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-408K4 is
  from the RP23-23 Mouse PAC Library
  constructed by the group of Pieter de Jong.
  For further details see http://www.chori.org/bacpac/home.htm
  VECTOR: pBACe3.6
  This sequence was finished as follows unless otherwise noted: all
  regions were either double-stranded or sequenced with an alternate
  chemistry or covered by high quality data (i.e., phred quality >=
  30); an attempt was made to resolve all sequencing problems, such
  as compressions and repeats; all regions were covered by at least
  one plasmid subclone or more than one M13 subclone; and the
  assembly was confirmed by restriction digest, except on the rare
  occasion of the clone being a YAC.
  Sequence from the Mouse Genome Sequencing Consortium whole genome

```

shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

FEATURES

Query Match	Best Local Similarity	79.9%;	Score 656.6;	DB 10;	Length 201642;
Matches 870;	Conservative 0;	Mismatches 189;	Indels 30;	Gaps 7;	
QY	69	AAGCTTTTGAGATTACAAATGCGCAGGAAATCATCACTTTTAAAGTAATTCCTCTCGA	128		
DB	60334	AACTTACATTATATATGCGCAGGAAATGCTCTTTTAAATTAATCTTCTTCGGA	60275		
QY	129	GATGTGGAGTTGGGAAGTTCACCTATGAAACAGATATGTAACATAATAGTTTCAATACC	188		
DB	60274	GATGTGGAGTTGGCAGAGTTCCTCTATGAAACAGATATGTAACCAATAATTTGATTC	60215		
QY	189	CAGCTCTTCATACATAGTGTGGAAATTTTAAATAAAAGATTTGGAAGTGAATGACAT	248		
DB	60214	CAGCTCTTCCACAAATAGTGTGGAAATTTTAAATAAAGATCTGGAGTGGACGACAT	60155		
QY	249	TTTGTTACATGAGATTTGGGACACGGCAGGTTCAGGACGATCCGAAGCCTGAGGACA	308		
DB	60154	TTTGTTACATGAGATTTGGGACACAGCTGGCCAAAGACCGTCCGAAGCCTGAGGACG	60095		
QY	309	CCATTTTACAGAGGTTCTGACTGCTCCCTGCTTACTTTTAGTGTGCGATGATTCACAAAGC	368		
DB	60094	CCATTTTACCGGGTCTGACTGCTGCTGCTTACATTTAGTGTGCGATGATTCACAGAGC	60035		
QY	369	TTCCAGAACTTAAGTAATCTGGAGAAAGAAATTCATATATATATGACAGATGTGAAGAGCCT	428		
DB	60034	TTCCAGAACTTGAGCAACTGGAGAAAGAGTTCATATATATGACAGATGTGAAGAGCCT	59975		
QY	429	GAGAGCTTTCCCTTTTGTGATCTGGGTAAACAAGATTGACATAAGCGAACGGCAGGTGCT	488		
DB	59974	GAAAGCTTTCCCTTTTGTGATTTGGCAACAAGCTGACATATAAGAACGGCAGGTGCT	59915		
QY	489	ACAGAGAAAGCCCAAGCTTTGGTCAGGGCAACACGGCGACTATCCTTATTTTGAACAAGT	548		
DB	59914	ACAGAGAAAGCCCAAGCTTTGGTCAGGGCAACACGGCGACTATCCTTATTTTGAACAAGT	59855		
QY	549	GCAAAAGATGCCACAATGTGGCAGCAGCCTTTTGAGGAGCGGTTGCAAGAGTTCCTTCT	608		
DB	59854	GCAAAAGATGCCACAATGTGCTGCTGCTTTGAGGAGCGAGTTGCAAGAAATTCCTGGCC	59795		
QY	609	ACCGAGGATAGGTTCAGATCATTTGATTCAGACAGACACAGTCAATCTTCCACGAAAGCCC	668		
DB	59794	ACCGAGGATAGGTTCAGACACACCTGATTCAGACAGACACAGTCAATCTGCAACGAAAGCCC	59735		
QY	669	AAGCCTTAGCTCATCTTCTGCTTTGATTTG-----TTAGATTGTTGATTCATTTCAACCAAC	722		
DB	59734	AAGCCTTAGCTCATCTTCTGCTTTGATTTGAGGATAGTTGTCATTTTCAACCAAC	59675		
QY	723	TACACATATACAAAAATCAACATGGGATGGAGAGAGAAATTAGCGTTTCAGCAGTG	782		
DB	59674	ACACACACACACAC-----ACAGAAAGGTGGAGAAACAAATTAGTATGCAAGAGATT	59622		
QY	783	TATCATCTACTAATA-AAATTAACCTAATGTTGCTTCAATTAGTGTGGTGGGAGAGGG	841		
DB	59621	CATTACTATATAAAATTCAGTTAAACCATATTGTTGGCTCATTTAGTGGTGGGAGAGGG	59562		
QY	842	ACACATCCACTCTTGGAGAAATATATTTACTCAATAATGGCACCTTACATTTATAAATTG	901		
DB	59561	ACACTCACTC---TGAGAGAAATCTATTTACTCAGTAATGGCAGTTTCACTTATAAATTG	59505		

Qy	902	TAACAGTTCTCTAATAAGCTTTTCTTTTAAATATGTAAGTTGCAGAGCTAATAAATG	961
Db	59504	TAATTGTTGCTAAAT---GTTTCTTTAAATTAACATTT-----AGTGTCTAATAAGAT	59455
Qy	962	AAATGACCAAGACTTTTAAATATAATAAAATAAGAAACTTGCATATTTCTAGAAGT---	1018
Db	59454	GACCAAGAAGTCACTTACTGTAAATCAGACGACAACTTGGCTAATTTAGAAGTACACT	59395
Qy	1019	TACTTGGATTTTTTCTGGGAAAAATGGAGAACTACTTTTTATGTGTAATGTTTTATGC	1078
Db	59394	TAGATTGTTTTCACCCACCAAGGAAATGGAAAATTACTTTATGTGTATATTTTATGT	59335
Qy	1079	AATTAGCATTTGTAATCTTGGTTTCAGGGAATACTTTTCCTAAAGCAATAATGTTAGATATT	1138
Db	59334	AATTAGCATTTGTAATCTTGGTCTTAGAAAGTAAATCTTACAGCAGTACTATTAAAGATT	59275
Qy	1139	AAAGATTAA	1147
Db	59274	AAATCTAA	59266

POSTAL 15

RESULT 15	AL672047	184474 bp	DNA	linear	ROD 18-SEP-2002
LOCUS	AL672047				
DEFINITION	Mouse DNA sequence from clone RP23-316B4 on chromosome X, complete sequence.				
ACCESSION	AL672047				
VERSION	AL672047.9	GI:23093223			
KEYWORDS	HTG.				
SOURCE	Mus musculus	(house mouse)			
ORGANISM	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1. (bases 1 to 184474)				
AUTHORS	Leonardorn1ert.D.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-SEP-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk				
COMMENT	On Sep 17, 2002 this sequence version replaced gi:22265406.				

COMMENT

COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: <http://www.sanger.ac.uk>
 Contact: hunguerv@sanger.ac.uk
 On Sep 17, 2002, this sequence version replaced gr:226263406.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emn., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-316B4 is from the RPI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACe3.6.

FEATURES

```
location/Qualifiers
source
1..184474
/organism="Mus musculus"
/mol type="genomic DNA"
```

Db	177919	AAATTAACAATT-----AGTGCTAATAAGATGACCAGAACTGACTTAATCTGTAATTC	1779771
QY	989	AAATAGAAACATTGACTATCTTAGAAGTTATATCTTTGGATTTTTT-----CCTGGGAAAAATG	1044
Db	177972	AGCAACAACCTTGACTATTCTTAGAAGTACACTTAGATTGTTTTGACCCACAAGGAAATG	178031
QY	1045	GAGAACTACTTTTTTATATGTGTATGTTTATTATGCAATTAGCAATTGATTTCTTGGTTCAGG	1104
Db	178032	GAGAAATTAC-TTTTTATATGTGTATATTTTATTGTAAATTAGCAATTGATTTCTTTGGTCTAGG	178090
QY	1105	GAAATACTTTTCTTAAGCAATAATGTTAGATATTAAAGATTAAAAATCTAAATCTAAAAAAA	1164
Db	178091	AAAGTAAATTCCTACAGC-AGATAAAATCTTAATTAAGATTAAATTTAATGCAAAAAAAA	178149
QY	1165	AAAAAANA	1173
Db	178150	AAAAAAGA	178158

Search completed: February 10, 2004, 03:56:06
Job time : 4605 secs

Db	177919	AAATTAACACTT-----AGTCTAATAAGATGACCAAGCACTGACTTAACCTGTAATTC	177971
QY	989	AAATTAACAACTTGGACTATCTAGAAAGTTATACCTTGGATTTTTT-----CCTGGGAAAATG	1044
Db	177972	AGACACACACCTTGGACTATCTAGAAAGTTATACCTTGGATTTTTTGGACCCCAAGGAAATG	178031
QY	1045	GAGAACTTACTTTTTTATATGTATGTTTATGCAATTTAGCATTTGATTTCTTGGTTTCAGG	1104
Db	178032	GAGAACTTAC-TTTTATATGTATATATTTTATGTAATTTAGCATTTGATTTCTTGGTTTCAGG	178090
QY	1105	GAAATCTTTTCCPAAAGCAATAATGTAGATATTAAGATTTAAATCTAATGTAAAAAAA	1164
Db	178091	AAAGTAAATTCCTACAGC-AGATAAAAAATCTAATAAGATTTAAATTTAATGCAAAAAA	178149
QY	1165	AAAAAAA 1173	
Db	178150	AAAAAAGA 178158	

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 9, 2004, 23:51:41 ; Search time 390 seconds
(without alignments)
8132.922 Million cell updates/sec

Title: US-09-988-974-6

Perfect score: 1175

Sequence: 1 CACTGTGATGAACACTTTT.....GTAAAAA.....AAAAA 1175

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1175	100.0	1175	20 AAX04755	CDNA encoding a Rab
2	1175	100.0	1175	24 ABQ75524	Human Rab protein
3	1160.2	98.7	1372	24 AB235377	Human gene express
4	1100	93.6	1969	22 ABA45112	Human breast cell
5	1100	93.6	1969	22 AKK03817	Human brain expres
6	1100	93.6	1969	22 AAI13885	Probe #3818 for ge
7	1100	93.6	1969	22 AAI03747	Probe #3738 used t
8	1100	93.6	1969	24 ABS03847	Human genome-deriv

9	1100	93.6	6058	22 AAL02805	Human reproductive
10	632	53.8	632	22 ABA50249	Human breast cell
11	632	53.8	632	22 AAK16569	Human brain expres
12	632	53.8	632	22 AAI23103	Probe #13036 for g
13	632	53.8	632	22 AAI08751	Probe #8742 used t
14	632	53.8	632	24 ABS16384	Human genome-deriv
15	604.4	51.4	606	25 ACA56484	Human signalling p
16	439.6	37.4	487	19 AAV61344	Extended CDNA seq
17	439.6	37.4	487	19 AAV58676	Prostate tumour sp
18	439.6	37.4	487	22 AAS63847	Human prostate cDN
19	439.6	37.4	487	22 AAS10198	Human prostate tum
20	439.6	37.4	487	22 AAH93555	Human prostate-spe
21	439.6	37.4	487	22 AAH84869	Human prostate-spe
22	439.6	37.4	487	22 AAH02620	Prostate tumour an
23	439.6	37.4	487	24 ABS71344	Prostate tumour tum
24	439.6	37.4	487	24 ABS58728	Prostate tumour CD
25	439.6	37.4	487	24 ABL95019	Human ID-4280 CDNA
26	439.6	37.4	487	25 ACA59456	Prostate cancer th
27	432.6	36.8	480	21 AAA06439	Human immunogenic
28	405.8	34.5	566	23 ABV59019	Human prostate exp
29	385.4	32.8	399	21 AAC30465	Human secreted pro
30	356	30.3	393	24 AB235026	Human gene express
31	333.2	28.4	1099	21 AA286894	Nucleoside sequenc
32	333.2	28.4	1114	22 AAS27473	CDNA encoding nove
33	333.2	28.4	1135	24 ABQ60937	CDNA encoding nove
34	328.4	27.9	1050	23 AAS91757	RAB9-like protein
35	322.4	27.4	1111	22 AAS27054	CDNA encoding nove
36	306.2	26.1	473	19 AAV61314	CDNA sequence of p
37	306.2	26.1	473	19 AAV58583	Prostate tumour sp
38	306.2	26.1	473	21 AAA06346	Human immunogenic
39	306.2	26.1	473	22 AAS63554	Human prostate cDN
40	306.2	26.1	473	22 AAS10105	Human prostate tum
41	306.2	26.1	473	22 AAH93462	Human prostate-spe
42	306.2	26.1	473	22 AAH84776	Human prostate-spe
43	306.2	26.1	473	22 AAH02527	Prostate tumour an
44	306.2	26.1	473	24 ABS71351	Human prostate tum
45	306.2	26.1	473	24 ABS58635	Prostate tumour CD

ALIGNMENTS

RESULT 1
AAX04755
ID AAX04755 standard; cDNA; 1175 BP.
XX AAX04755;
AC AAX04755 (first entry)
DT 08-APR-1999
XX cDNA encoding a Rab protein designated HRAB3-3.
DE Rab protein; HRAB3-3; Incyte clone 641412; vesicle targeting;
KW membrane transfer; protein processing; targeting; secretion;
KW abnormal neurotransmitter production; lysosomal storage disease;
KW choroideraemia; endocrine hypofunction; hyperfunction; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH CDS 87..692
FT /*tag= a
FT /product= HRAB33
XX US5869291-A.
XX 09-FEB-1999.
XX 27-DEC-1996; 96US-0773423.
XX 27-DEC-1996; 96US-0773423.
XX (INCY-) INCYTE PHARM INC.

PR 17-DEC-1998; 98US-0215887.
 XX (INCY-) INCYTE PHARM INC.
 PA Hillman JL, Goli SK, Bandman O;
 XX P-PSDB; ABP53078.
 DR WPI; 2002-655836/70.
 XX Novel Rab polypeptides, low molecular weight guanidine triphosphatases
 PT of Ras superfamily, useful for treating, preventing or diagnosing
 PT cancer, and immune system disorders e.g. ulcer, arthritis and Crohn's
 PT disease -
 XX Claim 5; Fig 3A-B; 40pp; English.
 PS The present sequence encodes human HRAB-3 which is a Rab protein (low
 XX molecular weight guanidine triphosphatases of the Ras superfamily).
 CC Rab proteins (i) have anti-neuritic, antidiabetic, immunosuppressive,
 CC anti-inflammatory, dermatological, neuroprotective, anti-rheumatoid,
 CC anti-ulcer, anti-arthritis, anti-HIV, cytostatic and endocrine activities,
 CC and can be used in gene therapy. (i) can be used for screening a compound
 CC for effectiveness as an agonist or antagonist, for screening a compound
 CC that specifically binds to (i) or modulates the activity of (i), and for
 CC preparing a polyclonal or monoclonal antibody by hybridoma technology.
 CC (i) can be used for diagnosing, treating or preventing immune system
 CC disorders (including autoimmune haemolytic anaemia, juvenile diabetes
 CC mellitus, lupus erythematosus, multiple sclerosis, myasthenia gravis,
 CC Crohn's disease, rheumatoid arthritis, AIDS or ulcerative colitis),
 CC cancer (including leukaemia, lymphoma, carcinomas of lung, penis, breast,
 CC pancreas, colon, stomach, small intestine, brain or prostate), diseases
 CC involving vesicle targeting, membrane transfer or fusion, or protein
 CC processing, targeting or secretion, and conditions characterised by
 CC abnormal levels of neurotransmitter production and release, lysosomal
 CC storage diseases, choroideraemia, endocrine hypofunction and
 CC hyperfunction.
 XX
 SQ Sequence 1175 BP; 387 A; 191 C; 244 G; 353 T; 0 other;
 Query Match 100.0%; Score 1175; DB 24; Length 1175;
 Best Local Similarity 100.0%; Pred. No. 5.9e-244;
 Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACTGTGATGAACACTTTCCCGTGTCTTTGAGTGCATCTTCTCAACACCCCTAGGAG 60
 DB 1 CACTGTGATGAACACTTTCCCGTGTCTTTGAGTGCATCTTCTCAACACCCCTAGGAG 60
 QY 61 GGTCTCTGAAGCTTTTGAGATTAACAATGGCAGGAAATCATCACTTTTAAAGTAATC 120
 DB 61 GGTCTCTGAAGCTTTTGAGATTAACAATGGCAGGAAATCATCACTTTTAAAGTAATC 120
 QY 121 TCCCTTGGAGATGGTGGAGTTGGAGAGTTCACTTATGAACAGATATGTAACATAAAGT 180
 DB 121 TCCCTTGGAGATGGTGGAGTTGGAGAGTTCACTTATGAACAGATATGTAACATAAAGT 180
 QY 181 TTGATACCCAGCTCTTCCATACAAATAGTGTGGAATTTTAAATGAAGATTGGAGTGG 240
 DB 181 TTGATACCCAGCTCTTCCATACAAATAGTGTGGAATTTTAAATGAAGATTGGAGTGG 240
 QY 241 ATGGACATTTTGTACATGAGATTTGGGACACGGCAGGTGAGGAGGATTCGGAAGCC 300
 DB 241 ATGGACATTTTGTACATGAGATTTGGGACACGGCAGGTGAGGAGGATTCGGAAGCC 300
 QY 301 TCAGGACACCATTTTACAGAGTTCTGACTGTGCTGCTACTTTTATGTCGATGATT 360
 DB 301 TCAGGACACCATTTTACAGAGTTCTGACTGTGCTGCTACTTTTATGTCGATGATT 360
 QY 361 CACAAAGCTTCCGAACCTTAAGTAACTGGAAGAAAGATTCATATATTATGCAATGTGA 420
 DB 361 CACAAAGCTTCCGAACCTTAAGTAACTGGAAGAAAGATTCATATATTATGCAATGTGA 420
 QY 421 AAGACCTGAGAGCTTCTCTTTTGTGATTTCTGGGTAAACAGATTGACATAGCGAAGCCG 480

DB 421 AAGACCTGAGAGCTTCTCTTTTGTGATTTCTGGGTAAACAGATTGACATAGCGAAGCCG 480
 QY 481 AGGTGTCTACAGAAAGACCCCAAGCTTGTGTCAGGGAACAACGGGCACTATCTTTATTTT 540
 DB 481 AGGTGTCTACAGAAAGACCCCAAGCTTGTGTCAGGGAACAACGGGCACTATCTTTATTTT 540
 QY 541 AAACAAGTGAAGAGATGCGACCAAAATCTGCGACAGCCTTTGAGGAAGCGGTTCGAAGAG 600
 DB 541 AAACAAGTGAAGAGATGCGACCAAAATCTGCGACAGCCTTTGAGGAAGCGGTTCGAAGAG 600
 QY 601 TTTTGTCTACCGAGATAGTGCAGATCATTTGATTTACAGACAGACACAGTCAATCTTCACC 660
 DB 601 TTTTGTCTACCGAGATAGTGCAGATCATTTGATTTACAGACAGACACAGTCAATCTTCACC 660
 QY 661 GAAAGCCCAAGCCTAGCTCATCTGCTGTTGATTTGATTTGATTTGATTTGATTTGATTTA 720
 DB 661 GAAAGCCCAAGCCTAGCTCATCTGCTGTTGATTTGATTTGATTTGATTTGATTTGATTTA 720
 QY 721 ACTCACACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCCGTTTGCAGCAG 780
 DB 721 ACTCACACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCCGTTTGCAGCAG 780
 QY 781 TGTATCATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
 DB 781 TGTATCATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 840
 QY 841 GACATATCCACTCTTGGAGCAATATATTTACTCAATAATGGCACCCTTACATTTTATAAAT 900
 DB 841 GACATATCCACTCTTGGAGCAATATATTTACTCAATAATGGCACCCTTACATTTTATAAAT 900
 QY 901 GTAACAGTGTCTAATAACCGTTCTTTTAAATTAATAATGTAAGTTGCAGAGCTAATAAT 960
 DB 901 GTAACAGTGTCTAATAACCGTTCTTTTAAATTAATAATGTAAGTTGCAGAGCTAATAAT 960
 QY 961 GAAATGACCAAGACTTTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1020
 DB 961 GAAATGACCAAGACTTTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1020
 QY 1021 CTTGGATTTTCTCGGAAATGGAGAACTACTTTTATATGTTATGTTATGTTATGTTATGCAA 1080
 DB 1021 CTTGGATTTTCTCGGAAATGGAGAACTACTTTTATATGTTATGTTATGTTATGTTATGCAA 1080
 QY 1081 TTAGCATTTATTTCTTGGTTCAGGGAATACTTTCTTAAGCAATAATGTTAGATATTA 1140
 DB 1081 TTAGCATTTATTTCTTGGTTCAGGGAATACTTTCTTAAGCAATAATGTTAGATATTA 1140
 QY 1141 AGATTAAATCTAATGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1175
 DB 1141 AGATTAAATCTAATGTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1175

RESULT 3
 ABZ35377
 ID ABZ35377 standard; cDNA; 1372 BP.
 XX
 AC ABZ35377;
 XX
 DT 05-FEB-2003 (first entry)
 DE Human gene expression profile polynucleotide SEQ ID NO 488.
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 OS Homo sapiens.
 XX
 PN W0200274979-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 20-MAR-2002; 2002WO-US08456.

PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 1; SEQ ID NO 3807; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and Bt 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1969 BP; 624 A; 330 C; 408 G; 607 T; 0 other;

Query Match 93.6%; Score 1100; DB 22; Length 1969;
Best Local Similarity 100.0%; Pred. No. 9.8e-228;
Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

59 AGGGTCTTGAAGCTTTGAGATTAACAATGGGAGGAAATCATCACTTTTAAAGTAAT 118
311 AGGGTCTTGAAGCTTTGAGATTAACAATGGGAGGAAATCATCACTTTTAAAGTAAT 370
119 TCTCCTTGGAGATGGTGGAGTGGGAGAGTTCACCTTATGAACAGATATGTAACATAAA 178
371 TCTCCTTGGAGATGGTGGAGTGGGAGAGTTCACCTTATGAACAGATATGTAACATAAA 430
179 GTTTGATACCCAGCTTCCATACATAGTGTGGAAATTTTAAATAAGATTTGGAAGT 238
431 GTTTGATACCCAGCTTCCATACATAGTGTGGAAATTTTAAATAAGATTTGGAAGT 490
239 GGATGGACATTTTGTACCATGAGATTTGGGACACGGCAGGTCAGGAGCTTCCGAAG 298
491 GGATGGACATTTTGTACCATGAGATTTGGGACACGGCAGGTCAGGAGCTTCCGAAG 550
299 CTTGAGGACACATTTTACAGAGTTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGA 358
551 CTTGAGGACACATTTTACAGAGTTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGA 610
359 TTCACAAAGCTTCCAGAACTTATAGTAACCTGGAGAAAGAAATTCATATATATGCGATGT 418
611 TTCACAAAGCTTCCAGAACTTATAGTAACCTGGAGAAAGAAATTCATATATATGCGATGT 670
419 GAAAGAGCTGAGAGCTTTCCTTTTGTGATCTGGGTAAACAAGTTGACATAGGGAACG 478

Db 671 GAAAGAGCCTGAGAGCTTTCTCTTTTGTGATTTCTGGGTAAACAAGATTGACATAACGGAACG 730
QY 479 GCAGGTGTCTACAGAAGAGCCCAAGCTTGTGTGAGGAGCAACCGGCACTATCTTTATTT 538
Db 731 GCAGGTGTCTACAGAAGAGCCCAAGCTTGTGTGAGGAGCAACCGGCACTATCTTTATTT 790
QY 539 TGAACAAAGTGCACAAAGATGCCAABAATGTGGCAGCAGCCCTTTGAGGAAGCGGTTCGAAG 598
Db 791 TGAACAAAGTGCACAAAGATGCCAABAATGTGGCAGCAGCCCTTTGAGGAAGCGGTTCGAAG 850
QY 599 AGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACAGACAGACAGTCAATCTTCA 658
Db 851 AGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACAGACAGACAGTCAATCTTCA 910
QY 659 CCGAAGCCCAAGCCTAGCTCATCTTCTGTTGATTTAGATTTGATGATCTTCAAC 718
Db 911 CCGAAGCCCAAGCCTAGCTCATCTTCTGTTGATTTAGATTTGATGATCTTCAAC 970
QY 719 CAATCTCACATATACACAAATCAACATGGGATGGAGAAAGAAATAGCGTTTGCAGC 778
Db 971 CAATCTCACATATACACAAATCAACATGGGATGGAGAAAGAAATAGCGTTTGCAGC 1030
QY 779 AGTGTATCATCTACTATATAAAATTAACATTAATGTGCTGCTTCAATAGTTGGTGGAGAA 838
Db 1031 AGTGTATCATCTACTATATAAAATTAACATTAATGTGCTGCTTCAATAGTTGGTGGAGAA 1090
QY 839 GGGACACATCCACCTCTTGGAGGAATATATTTACTCAATATGCGACCTTACATTTATAAA 898
Db 1091 GGGACACATCCACCTCTTGGAGGAATATATTTACTCAATATGCGACCTTACATTTATAAA 1150
QY 899 TTGTAAACAGTTGTCTAATAACGTTTCTTTAATTTAATAATATGTAAGTTGCGAGCTAATAA 958
Db 1151 TTGTAAACAGTTGTCTAATAACGTTTCTTTAATTTAATAATATGTAAGTTGCGAGCTAATAA 1210
QY 959 ATGAATGACCAAGACTTTAATTAATAATAAATAAGAACTTACATTTCTAGAAGTTA 1018
Db 1211 ATGAATGACCAAGACTTTAATTAATAATAAATAAGAACTTACATTTCTAGAAGTTA 1270
QY 1019 TACTTGGATTTTCTCTGGGAAATGGAGAACTACTTTTATATGCTGATGTTTTTATGC 1078
Db 1271 TACTTGGATTTTCTCTGGGAAATGGAGAACTACTTTTATATGCTGATGTTTTTATGC 1330
QY 1079 AATTAGCATTGTATTTCTTGGTTCAGGGAATACCTTCTTAAGCAATATGTTAGATATT 1138
Db 1331 AATTAGCATTGTATTTCTTGGTTCAGGGAATACCTTCTTAAGCAATATGTTAGATATT 1390
QY 1139 AAAGATTAAAAATCTAATGTA 1158
Db 1391 AAAGATTAAAAATCTAATGTA 1410
RESULT 5
AAK03817
ID AAK03817 standard; DNA; 1969 Bp.
XX
AC AAK03817;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 3808.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
OS Homo sapiens.
XX
FN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX

PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 XX
 DR Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 3808; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 1969 BP; 624 A; 330 C; 408 G; 607 T; 0 other;
 Query Match 93.6%; Score 1100; DB 22; Length 1969;
 Best Local Similarity 100.0%; Pred. No. 9.8e-228;
 Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 59 AGGGTCTTGAAGCTTTGAGATTAACAATGGCAGGAAATCATCATTAAAGTAAT 118
 311 AGGGTCTTGAAGCTTTGAGATTAACAATGGCAGGAAATCATCATTAAAGTAAT 370
 119 TCTCCTTGAGATGGTGGAGTTGGGAGAGTTCACTTATGAACAGATATGTAACATAA 178
 371 TCTCCTTGAGATGGTGGAGTTGGGAGAGTTCACTTATGAACAGATATGTAACATAA 430
 179 GTTGTGATCCAGCTCTCCATACATAGGTGTGGAATTTTAAATAAGATTTGGAAGT 238
 431 GTTGTGATCCAGCTCTCCATACATAGGTGTGGAATTTTAAATAAGATTTGGAAGT 490
 239 GGATGACATTTTGTACATGACAGATTTGGGACACGCGAGTCAGGACGATTCGGAAG 298
 491 GGATGACATTTTGTACATGACAGATTTGGGACACGCGAGTCAGGACGATTCGGAAG 550
 299 CCTGAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGA 358
 551 CCTGAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGA 610
 359 TTCAAAAGCTTCCAGAACTTAAGTAACTGGGAAGAAAGAAATTCATATATATGCAAGATG 418
 611 TTCAAAAGCTTCCAGAACTTAAGTAACTGGGAAGAAAGAAATTCATATATATGCAAGATG 670
 419 GAAAGAGCTTGAGAGCTTTCTTTTGTGATCTGGTAACTGATGATGATGATGATGATG 478
 671 GAAAGAGCTTGAGAGCTTTCTTTTGTGATCTGGTAACTGATGATGATGATGATGATG 730
 479 GCAGGTGTCTACAGAAAGACCCAAAGCTTTGGTGCAGGACCAACGGCGACTATTCCTTATTT 538
 731 GCAGGTGTCTACAGAAAGACCCAAAGCTTTGGTGCAGGACCAACGGCGACTATTCCTTATTT 790
 539 TGAACAAGTGCAGAAAGATGCCAAGATGGTGCAGAGCTTTGAGGAGCGGTTTCGAAG 598
 791 TGAACAAGTGCAGAAAGATGCCAAGATGGTGCAGAGCTTTGAGGAGCGGTTTCGAAG 850
 599 AGTTCCTTGCTACCGAGGATAGTTCAGATCATTTGATTCAGACAGACAGATCAATCTTCA 658
 851 AGTTCCTTGCTACCGAGGATAGTTCAGATCATTTGATTCAGACAGACAGATCAATCTTCA 910

QY 659 CCGAAGCCCAAGCCTAGCTCATCTTCTGTTGATTTGTTAGATTTGTTAGTGCATTTCTAAC 718
 DB 911 CCGAAGCCCAAGCCTAGCTCATCTTCTGTTGATTTGTTAGATTTGTTAGTGCATTTCTAAC 970
 QY 719 CAACCTCACATATACACAAAATCAACATGGGATGGAGAGAGAAATTAGCGTTTGCGAGC 778
 DB 971 CAACCTCACATATACACAAAATCAACATGGGATGGAGAGAGAAATTAGCGTTTGCGAGC 1030
 QY 779 AGTGTATCATCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 838
 DB 1031 AGTGTATCATCTACTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1090
 QY 839 GGGACACATCCACTCTTGGAGGAATATATTTACTCTAATAATGCACCTTACATTTATAA 898
 DB 1091 GGGACACATCCACTCTTGGAGGAATATATTTACTCTAATAATGCACCTTACATTTATAA 1150
 QY 999 TTGTAAACAGTTGTCTAATAAACGTTTCTTTAAATTAATAATGTAAGTTGCAGACTAATAA 958
 DB 1151 TTGTAAACAGTTGTCTAATAAACGTTTCTTTAAATTAATAATGTAAGTTGCAGACTAATAA 1210
 QY 959 ATGAAATGACCAAGACTTTAATTATATAAATAAATAAATAAATAAATAAATAAATAA 1018
 DB 1211 ATGAAATGACCAAGACTTTAATTATATAAATAAATAAATAAATAAATAAATAAATAA 1270
 QY 1019 TACTTGGATTTTTCCTGGGAAAATGGAGAACTACTTTTATATGTAATGTAATGTAATG 1078
 DB 1271 TACTTGGATTTTTCCTGGGAAAATGGAGAACTACTTTTATATGTAATGTAATGTAATG 1330
 QY 1079 AATTAGCATTTGATCTTCTGTTCCAGGAAATACTTCTTAAAGCAATATGTTAGATATT 1138
 DB 1331 AATTAGCATTTGATCTTCTGTTCCAGGAAATACTTCTTAAAGCAATATGTTAGATATT 1390
 QY 1139 AAAGATTAAATCTAATGTA 1158
 DB 1391 AAAGATTAAATCTAATGTA 1410
 RESULT 6
 AA113885
 ID AA113885 standard; DNA; 1969 BP.
 XX
 AC AA113885;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Probe #3818 for gene expression analysis in human cervical cell sample.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 OS Homo sapiens.
 XX
 FN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00670.
 XX
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 FA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488901/53.
 XX
 PI Human genome-derived single exon nucleic acid probes useful for

CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1969 BP; 624 A; 330 C; 408 G; 607 T; 0 other;
 Query Match 93.6%; Score 1100; DB 22; Length 1969;
 Best Local Similarity 100.0%; Pred. No. 9.8e-228;
 Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 AGGGTTCTTGAAGCTTTTGGAGTAACCAATGCGAGAAATCATCCTTTTAAAGTAAT 118
 DB 311 AGGGTTCTTGAAGCTTTTGGAGTAACCAATGCGAGAAATCATCCTTTTAAAGTAAT 370
 QY 119 TCTCCTGGAGATGCTGGAGTTGGAAGAGTTCATTATGAACAGATATGTAACATAA 178
 DB 371 TCTCCTGGAGATGCTGGAGTTGGAAGAGTTCATTATGAACAGATATGTAACATAA 430
 QY 179 GTTTCATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTAAATAAGATTTGGAAGT 238
 DB 431 GTTTCATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTAAATAAGATTTGGAAGT 490
 QY 239 GGATGGAATTTGTTTACATGACAGATTTGGACACGCGAGGTGAGGCGATTCGGAAG 298
 DB 491 GGATGGAATTTGTTTACATGACAGATTTGGACACGCGAGGTGAGGCGATTCGGAAG 550
 QY 299 CTTGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGCGATGA 358
 DB 551 CTTGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGCGATGA 610
 QY 359 TTCAAAAGCTTCCAGAACTTAAGTAATCGAAGAAAGATTCATATATATGAGATGT 418
 DB 611 TTCAAAAGCTTCCAGAACTTAAGTAATCGAAGAAAGATTCATATATATGAGATGT 670
 QY 419 GAAAGAGCTGAGAGCTTCTTTTGTGATCTGGTAAACAGATTCACATAAGCGAAG 478
 DB 671 GAAAGAGCTGAGAGCTTCTTTTGTGATCTGGTAAACAGATTCACATAAGCGAAG 730
 QY 479 GCAGGTGTCTACAGAAAGCCCAAGCTTGTGTGAGGACCAACGCGCATCTCTTATTT 538
 DB 731 GCAGGTGTCTACAGAAAGCCCAAGCTTGTGTGAGGACCAACGCGCATCTCTTATTT 790
 QY 539 TGAACAAGTGCAGAGATGCCAATAATGTGGCAGCAGCCTTTGAGAGCGGTTGGAAG 598
 DB 791 TGAACAAGTGCAGAGATGCCAATAATGTGGCAGCAGCCTTTGAGAGCGGTTGGAAG 850
 QY 599 AGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACACACAGTCATCTTCA 658
 DB 851 AGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACACACAGTCATCTTCA 910
 QY 659 CCGAAAGCCCAAGCTTAGCTCATCTTGTGTTGATTTAGATTTGATGATCATCTTAAC 718
 DB 911 CCGAAAGCCCAAGCTTAGCTCATCTTGTGTTGATTTAGATTTGATGATCATCTTAAC 970
 QY 719 CAACTCACATATACACAAATCAACATGGGATGGAGAGAGAAATTAGGTTTGGAG 778
 DB 971 CAACTCACATATACACAAATCAACATGGGATGGAGAGAGAAATTAGGTTTGGAG 1030
 QY 779 AGTGTATCATCTACTAATAAATAAATAAATTAAGTGTGCTTCAATAGTTCGGTGGAGAA 838
 DB 1031 AGTGTATCATCTACTAATAAATAAATAAATTAAGTGTGCTTCAATAGTTCGGTGGAGAA 1090
 QY 839 GGGACATCCACTCTTGGAGAAATATATTTACTCAATAAATGGCAGCTTACATTATAA 898
 DB 1091 GGGACATCCACTCTTGGAGAAATATATTTACTCAATAAATGGCAGCTTACATTATAA 1150
 QY 899 TTGTAACAGTGTCTAATAAAGCTTTCTTTTAAATATATATGTAAGTTCAGAGCTAATAA 958
 DB 1151 TTGTAACAGTGTCTAATAAAGCTTTCTTTTAAATATATGTAAGTTCAGAGCTAATAA 1210
 QY 959 ATGAAATGCCAAGACTTTAATATATAAATAAATAAAGAACTTGACTTATCTTAGAGTTA 1018

DB 1211 ATGAATGACCAAGACTTTAATATATAAATAAATAAAGAACTTGACTTATCTAGAGTTA 1270
 QY 1019 TACTTGGATTTTCTCTGGAAAAATGGAGAACTATCTTTTATGTATGTTTATGCT 1078
 DB 1271 TACTTGGATTTTCTCTGGAAAAATGGAGAACTATCTTTTATGTATGTTTATGCT 1330
 QY 1079 AATTAGCATTTGTTCTGTTTCAGGAAATACTTTCTTAAAGCAATATGTTAGATATT 1138
 DB 1331 AATTAGCATTTGTTCTGTTTCAGGAAATACTTTCTTAAAGCAATATGTTAGATATT 1390
 QY 1139 AAGATTAAATCTAATGTA 1158
 DB 1391 AAGATTAAATCTAATGTA 1410
 RESULT 8
 ABS03847
 ID ABS03847 standard; DNA; 1969 BP.
 XX
 AC ABS03847;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 3838.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Rudrak syndrome; sarcoidosis; pulmonary haemorrhage;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PP 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 XX
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 1; SEQ ID No 3838; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic,
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and lysine membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1969 BP; 624 A; 330 C; 408 G; 607 T; 0 other;

Query Match 93.6%; Score 1100; DB 24; Length 1969;
 Best Local Similarity 100.0%; Pred. No. 9.8e-228;
 Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 AGGGTTCTTGAAGCTTTTGAGATTAACAATGGCAGGAAATCATCTTTTAAAGTAAT 118
 DB 311 AGGGTTCTTGAAGCTTTTGAGATTAACAATGGCAGGAAATCATCTTTTAAAGTAAT 370
 QY 119 TCTCCTTGGAGATGGTGGAGTTGGGAGAGTTTCACTTATGAACAGATATGTAACATAA 178
 DB 371 TCTCCTTGGAGATGGTGGAGTTGGGAGAGTTTCACTTATGAACAGATATGTAACATAA 430
 QY 179 GTTTGATACCGACTCTTCCATACAAATAGTGTGGAAATTTTAAATAAGATTGGAAT 238
 DB 431 GTTTGATACCGACTCTTCCATACAAATAGTGTGGAAATTTTAAATAAGATTGGAAT 490
 QY 239 GGATGGACATTTTGTACCATGCAGATTTGGGACACGCGAGGTCAGGACGATTCGAG 298
 DB 491 GGATGGACATTTTGTACCATGCAGATTTGGGACACGCGAGGTCAGGACGATTCGAG 550
 QY 299 CTTGAGGACACCACTTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTCGATGA 358
 DB 551 CTTGAGGACACCACTTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTCGATGA 610
 QY 359 TTCCAAAGCTTCCAGACTTAAGTAACTGAGAGAAATTCATATATATGTCAGATGT 418
 DB 611 TTCCAAAGCTTCCAGACTTAAGTAACTGAGAGAAATTCATATATATGTCAGATGT 670
 QY 419 GAAAGAGCTGAGAGCTTTCTCTTTTCTGATTTCTGGGTAAACAAGATTGACATAAGCGA 478
 DB 671 GAAAGAGCTGAGAGCTTTCTCTTTTCTGATTTCTGGGTAAACAAGATTGACATAAGCGA 730
 QY 479 GCAGGTGTCTACAGAGAGCCCAAGCTTGTGTCAGGAGCAACGGCGACTATCTTATTT 538
 DB 731 GCAGGTGTCTACAGAGAGCCCAAGCTTGTGTCAGGAGCAACGGCGACTATCTTATTT 790
 QY 539 TGAACAAGTGAACCAAGATGCCAATAATGTGGCAGCAGCTTTTGAGGAAGCGGTCGAAG 598
 DB 791 TGAACAAGTGAACCAAGATGCCAATAATGTGGCAGCAGCTTTTGAGGAAGCGGTCGAAG 850

QY 599 AGTTCTTGCTACCGAGGATAGGTCAGATCATTTGATTTCAGACAGACACAGCTCAATCTTCA 658
 DB 851 AGTTCTTGCTACCGAGGATAGGTCAGATCATTTGATTTCAGACAGACACAGCTCAATCTTCA 910
 QY 659 CCGAAAGCCCAAGCCTAGCTCATCTTCTGTTGATTGTTTGTAGTTGTTGATGCTTAAAC 718
 DB 911 CCGAAAGCCCAAGCCTAGCTCATCTTCTGTTGATTGTTTGTAGTTGTTGATGCTTAAAC 970
 QY 719 CAATCTCACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCGTTTGCAGC 778
 DB 971 CAATCTCACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCGTTTGCAGC 1030
 QY 779 AGTGTATCATCTACTATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 838
 DB 1031 AGTGTATCATCTACTATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1090
 QY 839 GGGACACATCCACTCTTGGAGGAATATATTACTCAATAATGGCACCTTACATTTATAAA 898
 DB 1091 GGGACACATCCACTCTTGGAGGAATATATTACTCAATAATGGCACCTTACATTTATAAA 1150
 QY 899 TTGTAACAGTTGTTCTAATAACGTTTCTTAAATTAATAATGTAAGTTGCAGAGCTAATAA 958
 DB 1151 TTGTAACAGTTGTTCTAATAACGTTTCTTAAATTAATAATGTAAGTTGCAGAGCTAATAA 1210
 QY 959 ATGAAATGACCAAGACTTTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1018
 DB 1211 ATGAAATGACCAAGACTTTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1270
 QY 1019 TACTTGGATTTTTCTCGGGAATGAGAACTACTTTTATATGTTGTTTATGCTTTTATGC 1078
 DB 1271 TACTTGGATTTTTCTCGGGAATGAGAACTACTTTTATATGTTGTTTATGCTTTTATGC 1330
 QY 1079 AATTAGCATTTGTTTCTTGGTTTCAGGGAATACTTCTCTAAAGCAATAATGTTAGATATT 1138
 DB 1331 AATTAGCATTTGTTTCTTGGTTTCAGGGAATACTTCTCTAAAGCAATAATGTTAGATATT 1390
 QY 1139 AAGATTAAATCTAATGTA 1158
 DB 1391 AAGATTAAATCTAATGTA 1410
 RESULT 9
 AAL02805
 ID AAL02805 standard; DNA; 6058 BP.
 AC AAL02805;
 DT 21-NOV-2001 (first entry)
 XX Human reproductive system related antigen DNA SEQ ID NO: 5493.
 DE Human; reproductive system related antigen; reproductive system disorder;
 KW cancer; gene therapy; ds.
 XX Homo sapiens.
 XX WO200155320-A2.
 XX 02-AUG-2001.
 XX 17-JAN-2001; 2001WO-US01339.
 XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226686.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 20-OCT-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 08-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 11-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 03-JAN-2001; 2001US-0259578.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Barash SC, Ruben SM;
 XX
 XX WPI; 2001-465570/50.
 XX
 XX Isolated nucleic acid molecule encoding a reproductive system antigen -
 XX is used in preventing, treating or ameliorating a medical condition -
 XX Disclosure; SEQ ID NO 5493; 1297pp + Sequence Listing; English.
 XX
 XX The present invention provides the protein and coding sequences of a
 XX number of human reproductive system related antigens. These can be used
 XX in the prevention and treatment of reproductive system disorders,
 XX including cancer. The present sequence is a genomic sequence encoding a
 XX protein of the invention.
 XX
 XX Sequence 6058 BP; 1764 A; 1068 C; 1313 G; 1913 T; 0 other;
 SQ

Query Match 93.8%; Score 1100; DB 22; Length 6058;

Best Local Similarity 100.0%; Pred. No. 1.2e-227;		Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	59	AGGTTCTTGAAGCTTTGAGATTAAACATGCGCAGGAAATCATCTTTTAAAGTAAT	118
Db	4905	AGGTTCTTGAAGCTTTGAGATTAAACATGCGCAGGAAATCATCTTTTAAAGTAAT	4964
Qy	119	TCTCTTGGAGATGGTGGAGTTGGAAGAGTTTCACTTATGAACAGATATCTAACTAATA	178
Db	4965	TCTCTTGGAGATGGTGGAGTTGGAAGAGTTTCACTTATGAACAGATATCTAACTAATA	5024
Qy	179	GTTTGATACCCAGCTCTTCCATACAAATAGTGTGGAAATTTTAAATAAGATTGGAAGT	238
Db	5025	GTTTGATACCCAGCTCTTCCATACAAATAGTGTGGAAATTTTAAATAAGATTGGAAGT	5084
Qy	239	GGATGGACATTTTGTACCATGAGATTGGGACAGCGAGGTCAGGAGGATCCGAG	298
Db	5085	GGATGGACATTTTGTACCATGAGATTGGGACAGCGAGGTCAGGAGGATCCGAG	5144
Qy	299	CCTGAGGACACCAATTTTACAGAGGTTCTGACTGCTGCTCTTACTTTTAGTGTGATGA	358
Db	5145	CCTGAGGACACCAATTTTACAGAGGTTCTGACTGCTGCTCTTACTTTTAGTGTGATGA	5204
Qy	359	TTCAAAAGCTTCAGAACTTAACTAGTAACTGGAAGAAAGATTATATATATGCGATGT	418
Db	5205	TTCAAAAGCTTCAGAACTTAACTAGTAACTGGAAGAAAGATTATATATATGCGATGT	5264
Qy	419	GAAGAGCCTCAGAGCTTTCTTTTGTGATCTCTGGGTAAACAAGATTGACATAAGCGAACG	478
Db	5265	GAAGAGCCTCAGAGCTTTCTTTTGTGATCTCTGGGTAAACAAGATTGACATAAGCGAACG	5324
Qy	479	GCAGGTCTTACAGAAAGCCCAAGCTTGGTGAGGACAAACGGGCACTATCTTATTT	538
Db	5325	GCAGGTCTTACAGAAAGCCCAAGCTTGGTGAGGACAAACGGGCACTATCTTATTT	5384
Qy	539	TGAACAAGTGCAAAAGATGCCAATGTGGCAGAGCTTTTGAGGAGCGGTCGAG	598
Db	5385	TGAACAAGTGCAAAAGATGCCAATGTGGCAGAGCTTTTGAGGAGCGGTCGAG	5444
Qy	599	AGTTCTTGTCACGAGGATAGGTGAGATCAATTTGATTCAGACAGACACAGTCAATCTTCA	658
Db	5445	AGTTCTTGTCACGAGGATAGGTGAGATCAATTTGATTCAGACAGACACAGTCAATCTTCA	5504
Qy	659	CCGAAAGCCCAAGCTTATCTTCTGCTGCTTCAATTTGATTTAGATTTGATTTTAC	718
Db	5505	CCGAAAGCCCAAGCTTATCTTCTGCTGCTTCAATTTGATTTAGATTTGATTTTAC	5564
Qy	719	CAACTCACATATACAAAATCAACATGGGATGGAGAGAAATTAGCGTTTGCAGC	778
Db	5565	CAACTCACATATACAAAATCAACATGGGATGGAGAGAAATTAGCGTTTGCAGC	5624
Qy	779	AGTGTATCATCTACTAATAAATAAATACTAATGTTGCTGCTTCAATTTAGTTGGGAGAA	838
Db	5625	AGTGTATCATCTACTAATAAATAAATACTAATGTTGCTGCTTCAATTTAGTTGGGAGAA	5684
Qy	839	GGACACATCCCTCTTGGAGGAATATATTACTCAATATGTCACCTTACATTTATATAA	898
Db	5685	GGACACATCCCTCTTGGAGGAATATATTACTCAATATGTCACCTTACATTTATATAA	5744
Qy	899	TTGTACAGTGTCTAATAACGTTTCTTTAATTTAAATATGTAAGTTGAGAGCTAATAA	958
Db	5745	TTGTACAGTGTCTAATAACGTTTCTTTAATTTAAATATGTAAGTTGAGAGCTAATAA	5804
Qy	959	ATGAATGACCAAGCTTATATATATAAATAAGAACTTACTATCTTCTAGAGTTA	1018
Db	5805	ATGAATGACCAAGCTTATATATATAAATAAGAACTTACTATCTTCTAGAGTTA	5864
Qy	1019	TACTTGGATTTTCTTGGGAAATGGAGAACTACTTTTATATGATGTATTTTATGTC	1078
Db	5865	TACTTGGATTTTCTTGGGAAATGGAGAACTACTTTTATATGATGTATTTTATGTC	5924
Qy	1079	AATTAGCATGTATCTTGGTTCAGGAAATACTTTCCCTAAAGCAATAATGTTAGATTT	1138

Db	5925	AATTAGCATGTATCTTGGTTCAGGAAATACTTTCCCTAAAGCAATAATGTTAGATTT	5984
Qy	1139	AAAGATTAAATCTAATGTA	1158
Db	5985	AAAGATTAAATCTAATGTA	6004
RESULT 10			
ABAS0249	ID ABAS0249 standard; DNA; 632 BP.		
XX	ABAS0249;		
XX	01-FEB-2002 (first entry)		
DE	Human breast cell single exon nucleic acid probe #8944.		
XX	Human; microarray; single exon probe; gene expression; breast;		
KW	disease; cancer; ss.		
OS	Homo sapiens.		
XX	WO200157271-A2.		
XX	09-AUG-2001.		
XX	30-JAN-2001; 2001WO-US00662.		
XX	04-FEB-2000; 2000US-0180312.		
PR	26-MAY-2000; 2000US-0207456.		
PR	30-JUN-2000; 2000US-0608408.		
PR	03-AUG-2000; 2000US-0632366.		
PR	21-SEP-2000; 2000US-0234687.		
PR	27-SEP-2000; 2000US-0236359.		
PR	04-OCT-2000; 2000GB-0024263.		
XX	(MOLE-) MOLECULAR DYNAMICS INC.		
PI	Penn SG, Hanzel DK, Chen W, Rank DR;		
XX	WPI; 2001-496933/54.		
XX	New spatially-addressable set of single exon nucleic acid probes,		
PT	useful for measuring gene expression in sample derived from human		
PT	breast, comprises number of single exon nucleic acid probes -		
XX	Claim 4; SEQ ID NO 8944; 327bp + sequence listing; English.		
XX	The invention relates to a spatially-addressable set of single exon		
CC	nucleic acid probes for measuring gene expression in a sample derived		
CC	from human breast and Br 474 cells. The method involves contacting		
CC	the probes with a collection of detectably labelled nucleic acids		
CC	derived from mRNA of human breast, and then measuring the label		
CC	bound to each probe of the microarray. The probes are useful for		
CC	verifying the expression of regions of genomic DNA predicted to		
CC	encode proteins. They are useful for gene discovery, and for		
CC	determining predisposition and/or prognosing breast disease. Gene		
CC	expression analysis is useful for assessing the toxicity of chemical		
CC	agents on cells. The microarray of this invention presents a far greater		
CC	diversity of probes for measuring gene expression, with far less bias		
CC	than expressed sequence tag microarrays. The method is suitable for		
CC	rapid production of functional information from genomic sequence. The		
CC	present sequence is a single exon nucleic acid probe of the invention.		
CC	Note: The sequence data for this patent did not form part of the		
CC	printed specification, but was obtained in electronic format directly		
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences.		
XX	Sequence 632 BP; 190 A; 115 C; 153 G; 174 T; 0 other;		

Query March 53.8%; Score 632; DB 22; Length 632;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GGTCTTGAAGCTTTTGAATTAACCAATGCGAGGAAATCATCATCTTTTAAAGTAATTC 120
 Db 1 GGTCTTGAAGCTTTTGAATTAACCAATGCGAGGAAATCATCATCTTTTAAAGTAATTC 60
 QY 121 TCCTTGGAGATGCTGAGTTGGGAAGTTCCTTATGAAACAGATATGTAATAAGT 180
 Db 61 TCCTTGGAGATGCTGAGTTGGGAAGTTCCTTATGAAACAGATATGTAATAAGT 120
 QY 181 TTGATACCCAGCTCTTCCATACATATAGGTGTGGATTTTAAATTAAGATTTGGAGTGG 240
 Db 121 TTGATACCCAGCTCTTCCATACATATAGGTGTGGATTTTAAATTAAGATTTGGAGTGG 180
 QY 241 ATGGACATTTTGTACCATGCGATTTGGGACACGCGAGGTGAGGAGCAATCCCAAGCC 300
 Db 181 ATGGACATTTTGTACCATGCGATTTGGGACACGCGAGGTGAGGAGCAATCCCAAGCC 240
 QY 301 TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAAATTAAGATTTGGAGTGG 360
 Db 241 TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAAATTAAGATTTGGAGTGG 300
 QY 361 CACAAAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 420
 Db 301 CACAAAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 360
 QY 421 AGAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 480
 Db 361 AGAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 420
 QY 481 AGGTGCTACAGAGAGCCCAAGCTTGTGAGTGTGGGACACGCGAGTATCTTATTTTG 540
 Db 421 AGGTGCTACAGAGAGCCCAAGCTTGTGAGTGTGGGACACGCGAGTATCTTATTTTG 480
 QY 541 AAACAAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGTTTGAAGAG 600
 Db 481 AAACAAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGTTTGAAGAG 540
 QY 601 TTCTTGTACCGAGGATAGTGCATATTTGATTCAGACACACAGTCAATCTTCACC 660
 Db 541 TTCTTGTACCGAGGATAGTGCATATTTGATTCAGACACACAGTCAATCTTCACC 600
 QY 661 GAAAGCCCAAGCTTAGCTCATCTTGTGTTGA 692
 Db 601 GAAAGCCCAAGCTTAGCTCATCTTGTGTTGA 632

RESULT 11

AAK16569
 ID AAK16569 standard; DNA; 632 BP.

AC AAK16569;
 XX

DT 05-NOV-2001 (first entry)
 XX

DE Human brain expressed single exon probe SEQ ID NO: 16560.
 XX

XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW Epilepsy; Cancer; ss.

XX Homo sapiens.
 OS

XX WO200157275-A2.
 PN

XX 09-AUG-2001.
 XX

XX 30-JAN-2001; 2001WO-US00667.
 PF

XX 04-FEB-2000; 2000US-0180312.
 PR

XX 26-MAY-2000; 2000US-0207456.
 PR

XX 30-JUN-2000; 2000US-0608408.
 PR

XX 03-AUG-2000; 2000US-0632366.
 PR

XX 21-SEP-2000; 2000US-0234687.
 PR

XX 27-SEP-2000; 2000US-0236359.
 PR

PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX MPI; 2001-483446/52.
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 PT
 PS Example 4; SEQ ID NO: 16560; 650pp + Sequence Listing; English.
 XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is one of the probes of the
 CC invention.
 XX
 SQ Sequence 632 BP; 190 A; 115 C; 153 G; 174 T; 0 other;

Query Match 53.8%; Score 632; DB 22; Length 632;
 Best Local Similarity 100.0%; Pred. No. 6.4e-127;
 Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GGTCTTGAAGCTTTTGAATTAACCAATGCGAGGAAATCATCATCTTTTAAAGTAATTC 120
 Db 1 GGTCTTGAAGCTTTTGAATTAACCAATGCGAGGAAATCATCATCTTTTAAAGTAATTC 60
 QY 121 TCCTTGGAGATGCTGAGTTGGGAAGTTCCTTATGAAACAGATATGTAATAAGT 180
 Db 61 TCCTTGGAGATGCTGAGTTGGGAAGTTCCTTATGAAACAGATATGTAATAAGT 120
 QY 181 TTGATACCCAGCTCTTCCATACATATAGGTGTGGATTTTAAATTAAGATTTGGAGTGG 240
 Db 121 TTGATACCCAGCTCTTCCATACATATAGGTGTGGATTTTAAATTAAGATTTGGAGTGG 180
 QY 241 ATGGACATTTTGTACCATGCGATTTGGGACACGCGAGGTGAGGAGCAATCCCAAGCC 300
 Db 181 ATGGACATTTTGTACCATGCGATTTGGGACACGCGAGGTGAGGAGCAATCCCAAGCC 240
 QY 301 TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAAATTAAGATTTGGAGTGG 360
 Db 241 TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAAATTAAGATTTGGAGTGG 300
 QY 361 CACAAAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 420
 Db 301 CACAAAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 360
 QY 421 AGAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 480
 Db 361 AGAGCTTCCAGAGCTTAAAGTAACTGGAAGAAAGATTTATATATGAGATGTA 420
 QY 481 AGGTGCTACAGAGAGCCCAAGCTTGTGAGTGTGGGACACGCGAGTATCTTATTTTG 540
 Db 421 AGGTGCTACAGAGAGCCCAAGCTTGTGAGTGTGGGACACGCGAGTATCTTATTTTG 480
 QY 541 AAACAAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGTTTGAAGAG 600
 Db 481 AAACAAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGTTTGAAGAG 540
 QY 601 TTCTTGTACCGAGGATAGTGCATATTTGATTCAGACACACAGTCAATCTTCACC 660
 Db 541 TTCTTGTACCGAGGATAGTGCATATTTGATTCAGACACACAGTCAATCTTCACC 600
 QY 661 GAAAGCCCAAGCTTAGCTCATCTTGTGTTGA 692
 Db 601 GAAAGCCCAAGCTTAGCTCATCTTGTGTTGA 632

RESULT 12

CC staging, monitoring and prognosing diseases of the human breast.
 CC particularly those diseases with polygenic aetiology. The diseases
 CC include: breast cancer, disorders of development, inflammatory diseases
 CC of the breast, fibrocystic changes, proliferative breast disease and
 CC non-carcinoma tumours.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pat_sequences.

XX
 SQ Sequence 632 BP; 190 A; 115 C; 153 G; 174 T; 0 other;
 Query Match 53.8%; Score 632; DB 22; Length 632;
 Best Local Similarity 100.0%; Pred. No. 6.4e-127;
 Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 GGTTCTTGAAGCTTTTGAATTAACATGCGAGGAAATCATCACTTTTAAAGTAATTC 120
 Db 1 GGTTCTTGAAGCTTTTGAATTAACATGCGAGGAAATCATCACTTTTAAAGTAATTC 60
 QY 121 TCCTTGGAGATGGTGGAGTTGGAGAGTTCCTTATGAACAGATATGTAATTAAGT 180
 Db TCCTTGGAGATGGTGGAGTTGGAGAGTTCCTTATGAACAGATATGTAATTAAGT 120
 QY 181 TTGATACCCAGCTCTTCCATACATAGTGTGGAATTTTAAAGTTTGGAGTGG 240
 Db TTGATACCCAGCTCTTCCATACATAGTGTGGAATTTTAAAGTTTGGAGTGG 180
 QY 241 ATGGACATTTTGTACCATTCAGATTTGGGACACGCGAGTTCAGGAGGATTTCCGAAGCC 300
 Db ATGGACATTTTGTACCATTCAGATTTGGGACACGCGAGTTCAGGAGGATTTCCGAAGCC 240
 QY 301 TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTCGATGTT 360
 Db TGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTGTGTCGATGTT 300
 QY 361 CACAAAGCTTCAGAACTTAAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 420
 Db CACAAAGCTTCAGAACTTAAAGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 360
 QY 421 AAGAGCTCAGAGCTTTTCTTTTGTGATTTCTGGTGAACAGATTCAGTGAAGCGG 480
 Db AAGAGCTCAGAGCTTTTCTTTTGTGATTTCTGGTGAACAGATTCAGTGAAGCGG 420
 QY 481 AGTGTCACAGAAAGCCCAAGCTTGTGTGAGGACACAGCGGACTATCTTATTTTG 540
 Db AGTGTCACAGAAAGCCCAAGCTTGTGTGAGGACACAGCGGACTATCTTATTTTG 480
 QY 541 AAACAGTGCAGAAAGATGCCAATGTGGCAGCGCTTTGAGGAGCGGTTCGAGAG 600
 Db AAACAGTGCAGAAAGATGCCAATGTGGCAGCGCTTTGAGGAGCGGTTCGAGAG 540
 QY 601 TTCTTGCTACGAGGATAGTTCAGATCTTTGATTCAGACAGACAGTCACTTCCACC 660
 Db TTCTTGCTACGAGGATAGTTCAGATCTTTGATTCAGACAGACAGTCACTTCCACC 541
 QY 661 GAAAGCCCAAGCTAGCTCATCTTGTGTTGA 692
 Db GAAAGCCCAAGCTAGCTCATCTTGTGTTGA 632

RESULT 14
 ABS16384
 ID ABS16384 standard; DNA; 632 BP.

XX ABS16384;
 AC ABS16384;
 DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe ORF from lung SEQ ID No 16375.
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 Hemansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 primary ciliary dyskinesia; pulmonary hypertension;
 hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.
 XX WO200106003-A2.
 XX 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 XX 04-FEB-2000; 2000US-180312P.
 XX 26-MAY-2000; 2000US-207456P.
 XX 30-JUN-2000; 2000US-0608408.
 XX 03-AUG-2000; 2000US-0632366.
 XX 21-SEP-2000; 2000US-234687P.
 XX 27-SEP-2000; 2000US-236359P.
 XX 04-OCT-2000; 2000GB-0024263.
 (MOLE-) MOLECULAR DYNAMICS INC.
 Penn SG, Hanzel DK, Chen W, Rank DR;
 WPI; 2002-114183/15.
 Spatially-addressable set of single exon nucleic acid probes, used to
 measure gene expression in human lung samples -
 Claim 4; SEQ ID No 16375; 634pp; English.

The invention relates to a spatially-addressable set of single exon
 nucleic acid probes for measuring gene expression in a sample derived
 from human lung comprising single exon nucleic acid probes having one of
 12614 nucleic acid sequences mentioned in the specification, or their
 complements or the 12387 open reading frames derived from the 12614
 probes. Also included are a microarray comprising the novel set of
 probes. The novel set of probes which hybridise at high stringency to a
 nucleic acid expressed in the human lung; measuring gene expression in a
 sample derived from human lung, comprising (a) contacting the array with
 a collection of detectably labeled nucleic acids derived from human lung
 mRNA, and (b) measuring the label detectably bound to each probe of
 the array; identifying exons in a eukaryotic genome, comprising
 (a) algorithmically predicting at least one exon from genomic sequences
 of the eukaryote; and (b) detecting specific hybridisation of detectably
 labeled nucleic acids from eukaryote lung mRNA, to a single exon probe.
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at

```
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 632 BP; 190 A; 115 C; 153 G; 174 T; 0 other;

Query Match      53.8%; Score 632; DB 24; Length 632;
Best Local Similarity 100.0%; Pred. No. 6.4e-127;
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 GGTCTCTGAAGCTTTTTCAGATTAACAATGGCAGGAAATCATCATCTTTTAAAGTAATTC 120
Db 1 GGTCTCTGAAGCTTTTTCAGATTAACAATGGCAGGAAATCATCATCTTTTAAAGTAATTC 60
Qy 121 TCCTTGGAGATGGTGGAGTGGGAGAGTTCACCTTATGAACAGATATGAACCTAATAAGT 180
Db 61 TCCTTGGAGATGGTGGAGTGGGAGAGTTCACCTTATGAACAGATATGAACCTAATAAGT 120
Qy 181 TTGATACCCAGCTCTTCCATACATAGCTGTGAATTTTAAATAAGATTGGAGTGG 240
Db 121 TTGATACCCAGCTCTTCCATACATAGCTGTGAATTTTAAATAAGATTGGAGTGG 180
Qy 241 ATGGACATTTTGTATACCATCAGATTTGGACACGGCAGGTCCAGGAGCGATTCCGAAGCC 300
Db 181 ATGGACATTTTGTATACCATCAGATTTGGACACGGCAGGTCCAGGAGCGATTCCGAAGCC 240
Qy 301 TGAGGACACCATTTTACAGAGTCTGACTGCTGCCCTTACTTTTAGTTCGATGAT 360
Db 241 TGAGGACACCATTTTACAGAGTCTGACTGCTGCCCTTACTTTTAGTTCGATGAT 300
Qy 361 CACAAGCTTCCAGAACTTAAGTAACCTGGAAGAAAGAAATTCATATATTATGCAGATGCA 420
Db 301 CACAAGCTTCCAGAACTTAAGTAACCTGGAAGAAAGAAATTCATATATTATGCAGATGCA 360
Qy 421 AAGAGCTGAGAGCTTTCCTTTTGTGATTCCTGGTAAACAGATTGACATGAAGCGAACGGC 480
Db 361 AAGAGCTGAGAGCTTTCCTTTTGTGATTCCTGGTAAACAGATTGACATGAAGCGAACGGC 420
Qy 481 AGGTGTCTACAGAAAGCCCAAGCTTGGTGGAGGACACAGCGGACTATCTTATTG 540
Db 421 AGGTGTCTACAGAAAGCCCAAGCTTGGTGGAGGACACAGCGGACTATCTTATTG 480
Qy 541 AAACAAGTGCAAAAGATGCGCAAAATGTGGCAGCAGCCTTTGAGGAAGCGTTGCAAGAG 600
Db 481 AAACAAGTGCAAAAGATGCGCAAAATGTGGCAGCAGCCTTTGAGGAAGCGTTGCAAGAG 540
Qy 601 TTCTTGTACCGAGGATAGTTCAGATCATTGATTCAGACAGACAGTCATCTTCACC 560
Db 541 TTCTTGTACCGAGGATAGTTCAGATCATTGATTCAGACAGACAGTCATCTTCACC 560
Qy 661 GAAAGCCCAAGCCTAGCTCATCTTGTGTTGA 692
Db 601 GAAAGCCCAAGCCTAGCTCATCTTGTGTTGA 632

RESULT 15
ACA56484
ID ACA56484 standard; cdna; 606 BP.
AC AC
XX AC
XX AC
DT 06-JUN-2003 (first entry)
XX
DE Human signalling pathway polynucleotide probe SEQ ID NO 1082.
XX
KW Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
PN US6500938-B1.
XX
PD 31-DEC-2002.
XX
```

```
PF 30-JAN-1998; 98US-0016434.
XX
PR 30-JAN-1998; 98US-0016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
DR WPI; 2003-352189/33.
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides
XX
PS Claim 1; SEQ ID NO 1082; 65pp; English.
XX
CC The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?DocID=06500938B1.
XX
SQ Sequence 606 BP; 182 A; 112 C; 147 G; 165 T; 0 other;

Query Match      51.4%; Score 604.4; DB 25; Length 606;
Best Local Similarity 99.8%; Pred. No. 5.7e-121;
Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 87 ATGGCAGGAAATCATCATCTTTTAAAGTAATTCCTCTGGAGATGGTGGAGTTGGGAAG 146
Db 1 ATGGCAGGAAATCTTCACITTTTAAAGTAATTCCTCTGGAGATGGTGGAGTTGGGAAG 60
Qy 147 AGTTCACCTATGAACAGATATGTAACCTANTAGTTGATACCCAGCTCTTCCATACATA 206
Db 61 AGTTCACCTATGAACAGATATGTAACCTANTAGTTGATACCCAGCTCTTCCATACATA 120
Qy 207 GGTGTGGAATTTTAAATAAAGATTTGGAAGTGGATGACATTTTGTATTACCATCAGATT 266
Db 121 GGTGTGGAATTTTAAATAAAGATTTGGAAGTGGATGACATTTTGTATTACCATCAGATT 180
Qy 267 TGGGACACGGCAGGTCCAGGAGCGATTCGGAAGCCTGAGGACACCATTTTACAGAGTTCT 326
Db 181 TGGGACACGGCAGGTCCAGGAGCGATTCGGAAGCCTGAGGACACCATTTTACAGAGTTCT 240
Qy 327 GACTGTGCTGCTTACTTTTGTGTCGATGATTCACAAAGCTTCCAGAACTTAAGTAAC 386
Db 241 GACTGTGCTGCTTACTTTTGTGTCGATGATTCACAAAGCTTCCAGAACTTAAGTAAC 300
Qy 387 TGGGAAGAAAGAAATTCATATATATGTCAGATGTGGAAGAGCCTGAGAGCTTTCCTTTGTG 446
Db 301 TGGGAAGAAAGAAATTCATATATATGTCAGATGTGGAAGAGCCTGAGAGCTTTCCTTTGTG 360
Qy 447 ATCTTGCTACAGATTTGACATGAAGCGGAGCGGTGTCTACAGAGAGGCCAGCT 506
Db 361 ATCTTGCTACAGATTTGACATGAAGCGGAGCGGTGTCTACAGAGAGGCCAGCT 420
```

QY	507	TGTTGAGGACAAACGGGACTATCCTTATTTTGAACAAGTGCAAAAGATGCCACAAT	566
Db	421	TGTTGAGGACAAACGGGACTATCCTTATTTTGAACAAGTGCAAAAGATGCCACAAT	480
QY	567	GTGGCAGCAGCCTTTGAGGAAGCGGTTGGAAGAGTTCTTTGCTACCGAGGATAGGTCAGAT	626
Db	481	GTGGCAGCAGCCTTTGAGGAAGCGGTTGGAAGAGTTCTTTGCTACCGAGGATAGGTCAGAT	540
QY	627	CATTGATTTCAGACACACAGTCATCTTCACCGAAGCCCAAGCCTAGCTCATCTTGC	686
Db	541	CATTGATTTCAGACACACAGTCATCTTCACCGAAGCCCAAGCCTAGCTCATCTTGC	600
QY	687	TGTTGA	692
Db	601	TGTTGA	606

Search completed: February 10, 2004, 02:39:12
Job time : 396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2004, 01:51:16 ; Search time 92 Seconds
(without alignments)
5637.230 Million cell updates/sec

Title: US-09-988-974-6

Perfect score: 1175

Sequence: 1 CACTGTGATGAACACTTTT.....GTAAAAA.....1175

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 113956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/prodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/prodata/2/ina/PCITUS_COMB.seq.*
- 6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1175	100.0	1175	2	US-08-773-423-6
2	604.4	51.4	606	4	US-09-016-434-1082
3	439.6	37.4	487	3	US-09-030-607-206
4	439.6	37.4	487	4	US-09-439-313-206
5	439.6	37.4	487	4	US-09-352-618A-206
6	439.6	37.4	487	4	US-09-232-149A-206
7	306.2	26.1	473	3	US-09-020-956-106
8	306.2	26.1	473	3	US-09-030-607-106
9	306.2	26.1	473	4	US-09-439-313-106
10	306.2	26.1	473	4	US-09-352-618A-106
11	306.2	26.1	473	4	US-09-232-149A-106
12	181.8	15.5	624	4	US-09-016-434-1083
13	98.8	8.4	970	3	US-08-888-077A-28
14	81.2	6.9	723	4	US-09-016-434-1422
15	69.6	5.9	944	4	US-09-016-434-1060
16	65.8	5.6	1069	4	US-09-620-312D-646
17	63.8	5.4	1533	4	US-09-075-454-11
18	63.6	5.4	1172	4	US-09-075-454-8
19	63.6	5.4	1255	2	US-08-766-551-6
20	63.6	5.4	1886	4	US-09-620-312D-647
21	62.8	5.3	925	2	US-08-916-901-4
22	62.8	5.3	925	4	US-09-154-602-4
23	60.2	5.1	286	4	US-09-313-294A-6296
24	60.2	5.1	651	4	US-09-016-434-1255
25	59.8	5.1	875	4	US-09-669-751-150
26	59.4	5.1	875	4	US-09-075-454-10
27	59.4	5.1	1106	4	US-09-620-312D-959

28	59.4	5.1	2612	4	US-09-484-970B-142	Sequence 142, Appl
29	58.6	5.0	730	4	US-09-300-958A-16	Sequence 16, Appl
30	58	4.9	639	4	US-09-399-313-66	Sequence 66, Appl
31	55.4	4.7	1631	4	US-09-620-312D-587	Sequence 587, Appl
32	55	4.7	472	4	US-09-325-932A-30	Sequence 30, Appl
33	53.8	4.6	833	4	US-09-620-312D-426	Sequence 426, Appl
34	53.6	4.6	803	4	US-09-075-454-13	Sequence 13, Appl
35	53.2	4.5	1340	2	US-08-824-873-2	Sequence 2, Appl
36	53.2	4.5	1340	3	US-09-198-184-2	Sequence 2, Appl
37	53.2	4.5	1407	4	US-09-433-314-1	Sequence 1, Appl
38	52.8	4.5	847	2	US-08-773-423-4	Sequence 4, Appl
39	51.2	4.4	1102	4	US-09-620-312D-828	Sequence 828, Appl
40	51.2	4.4	1127	4	US-09-391-741A-1	Sequence 1, Appl
41	51.2	4.4	1127	4	US-09-391-741A-25	Sequence 25, Appl
42	51	4.3	241	4	US-09-016-434-813	Sequence 813, Appl
43	51	4.3	465	4	US-09-325-932A-34	Sequence 34, Appl
44	50	4.3	719	2	US-08-766-551-2	Sequence 2, Appl
45	50	4.3	719	4	US-09-016-434-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1
US-08-773-423-6
; Sequence 6, Application US/08773423
; Patent No. 5869291
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/773,423
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0183 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-08-773-423-6

Query Match 100.0%; Score 1175; DB 2; Length 1175;
Best Local Similarity 100.0%; Pred. No. 6e-286;

Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACTGTGATGAACACTTTTCCCGTGTGAGTGCATCTTCTCAACACCCCTAGGAG 60
 Db 1 CACTGTGATGAACACTTTTCCCGTGTGAGTGCATCTTCTCAACACCCCTAGGAG 60

QY 61 GGTCTCTGAGCTTTTGAATTAACAATGGCGAGGAATATCATCTTTTAAAGTAAATTC 120
 Db 61 GGTCTCTGAGCTTTTGAATTAACAATGGCGAGGAATATCATCTTTTAAAGTAAATTC 120

QY 121 TCCTTGGAGATGGTGGAGTTGGGAAGATTCACATATGAACAGATATGTAACTAATAGT 180
 Db 121 TCCTTGGAGATGGTGGAGTTGGGAAGATTCACATATGAACAGATATGTAACTAATAGT 180

QY 181 TTGATACCCAGCTCTTCCATCAATAGGTGTGGATTTTAAATAAAGATTTGGAAGTGG 240
 Db 181 TTGATACCCAGCTCTTCCATCAATAGGTGTGGATTTTAAATAAAGATTTGGAAGTGG 240

QY 241 ATGGACATTTTGTACCATGTCAGATTTGGGACACGGCAGGTTCAGGAGCGATTCGGAAGCC 300
 Db 241 ATGGACATTTTGTACCATGTCAGATTTGGGACACGGCAGGTTCAGGAGCGATTCGGAAGCC 300

QY 301 TGAGACACCATTTTACAGAGTCTGACTGCTGCTTACTTTTAGTGTGATGATTT 360
 Db 301 TGAGACACCATTTTACAGAGTCTGACTGCTGCTTACTTTTAGTGTGATGATTT 360

QY 361 CACAAAGCTTCAGAACTTAAAGTAACTGGAAGAAAGATTCATATATTATCAGATGTGA 420
 Db 361 CACAAAGCTTCAGAACTTAAAGTAACTGGAAGAAAGATTCATATATTATCAGATGTGA 420

QY 421 AAGAGCTTGAGAGCTTTCTTTTGTGATCTGGGTAAACAAGATTTGACATAGCGAAGCGC 480
 Db 421 AAGAGCTTGAGAGCTTTCTTTTGTGATCTGGGTAAACAAGATTTGACATAGCGAAGCGC 480

QY 481 AGGTGCTTACAGAAAGCCCAAGCTTGTGTCAGGACCAACGGGACTATCTTTATTG 540
 Db 481 AGGTGCTTACAGAAAGCCCAAGCTTGTGTCAGGACCAACGGGACTATCTTTATTG 540

QY 541 AAACAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGGTTCGAAGAG 600
 Db 541 AAACAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCTTTGAGGAAGCGGTTCGAAGAG 600

QY 601 TTCTTGTACAGAGATAGGTGAGATCATTTGATTCACAGACACAGTCAATCTTCAAC 660
 Db 601 TTCTTGTACAGAGATAGGTGAGATCATTTGATTCACAGACACAGTCAATCTTCAAC 660

QY 661 GAAAGCCCAAGCTAGCTCATCTTCTGCTGTGATTTAGATTTGTTGATGCAATCTTAACCA 720
 Db 661 GAAAGCCCAAGCTAGCTCATCTTCTGCTGTGATTTAGATTTGTTGATGCAATCTTAACCA 720

QY 721 ACTCACATATACACAAATCAACATGGGATGGAGAGAGATTTAGCGTTTCAGCAG 780
 Db 721 ACTCACATATACACAAATCAACATGGGATGGAGAGAGATTTAGCGTTTCAGCAG 780

QY 781 TGTATCATCTAATAAATAAATAAATAATGCTGCTTCAATTTAGTTGGGAGAGG 840
 Db 781 TGTATCATCTAATAAATAAATAAATAATGCTGCTTCAATTTAGTTGGGAGAGG 840

QY 841 GACACATCCACTCTTGGAGGAATATATTTTCAATATGCAATGCACTTCAATTTAATTT 900
 Db 841 GACACATCCACTCTTGGAGGAATATATTTTCAATATGCAATGCACTTCAATTTAATTT 900

QY 901 GTAACAGTTGTCTAATAAGCTTTCTTTTAAATATGATGATGGAGAGCTAATAAT 960
 Db 901 GTAACAGTTGTCTAATAAGCTTTCTTTTAAATATGATGATGGAGAGCTAATAAT 960

QY 961 GAAATGACCAAGACTTTAATAATAAATAAATAAAGAACTTGACTATTTCTAGAGTTATA 1020
 Db 961 GAAATGACCAAGACTTTAATAATAAATAAATAAAGAACTTGACTATTTCTAGAGTTATA 1020

QY 1021 CTTGGATTTTTTCTGGGAAATGAGAGCTACTTTTATATGCTATGCTTTTATGCA 1080
 Db 1021 CTTGGATTTTTTCTGGGAAATGAGAGCTACTTTTATATGCTATGCTTTTATGCA 1080

QY 1081 TTAGCATTTGATTTCTTGGTTTCTAGGAAATACTTTCTTAAGCAATAATGTTAGATATAA 1140
 Db 1081 TTAGCATTTGATTTCTTGGTTTCTAGGAAATACTTTCTTAAGCAATAATGTTAGATATAA 1140

QY 1141 AGATTAAATCTAATGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1175
 Db 1141 AGATTAAATCTAATGTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1175

RESULT 2
 US-09-016-434-1082
 ; Sequence 1082, Application US/09016434
 ; Patent No. 6500938
 ; GENERAL INFORMATION:
 ; APPLICANT: Janice Au-Young
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
 ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
 ; NUMBER OF SEQUENCES: 1490
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94304
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/016,434
 ; FILING DATE: HEREWITH
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Zeller, Karen J.
 ; REGISTRATION NUMBER: 37,071
 ; REFERENCE/DOCKET NUMBER: PA-0002 US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (650) 855-0555
 ; TELEFAX: (650) 845-4166
 ; INFORMATION FOR SEQ ID NO: 1082:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 606 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; LIBRARY: GENBANK
 ; CLONE: g1174146
 ; US-09-016-434-1082

Query Match 51.4%; Score 604.4; DB 4; Length 606;
 Best Local Similarity 99.8%; Pred. No. 9.9e-143;
 Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 87 ATGGCAGGAAATCATCACTTTTAAAGTAATTCCTTGGAGATGGTGGAGTTGGGAAG 146
 Db 1 ATGGCAGGAAATCTTCATCTTTTAAAGTAATTCCTTGGAGATGGTGGAGTTGGGAAG 60

QY 147 AGTTCACTTATGACAGATATGTAATAAGTTTGAATGATGATGATGATGATGATGATGAT 206
 Db 61 AGTTCACTTATGACAGATATGTAATAAGTTTGAATGATGATGATGATGATGATGATGAT 120

QY 207 GGTGTGGAATTTTAAATAAAGATTTTGAAGTGGATGGACATTTTGTACCATGCAGATT 266
 Db 121 GGTGTGGAATTTTAAATAAAGATTTTGAAGTGGATGGACATTTTGTACCATGCAGATT 180

QY 267 TGGGACGCGCAGGTGAGGAGCGATCCGAGGCTGAGGACACCACTTTTACAGAGTTCT 326
Db 181 TGGGACGCGCAGGTGAGGAGCGATCCGAGGCTGAGGACACCACTTTTACAGAGTTCT 240
QY 327 GACTGCTGCTGCTTACTTTTAGTGTGATGATTCACAAAGCTTCAGAACTTTAAGTAAC 386
Db 241 GACTGCTGCTGCTTACTTTTAGTGTGATGATTCACAAAGCTTCAGAACTTTAAGTAAC 300
QY 387 TGAAGAAGAAATTCATATATTATGAGATGTAAGAGCTGAGAGCTTCTCTTTTGTG 446
Db 301 TGAAGAAGAAATTCATATATTATGAGATGTAAGAGCTGAGAGCTTCTCTTTTGTG 360
QY 447 ATTCTGGGTAAACAAGATTGACATAAGCGACGCGGTCTACAGAAAGCCCAAGCT 506
Db 361 ATTCTGGGTAAACAAGATTGACATAAGCGACGCGGTCTACAGAAAGCCCAAGCT 420
QY 507 TGTGACGGGACAAAGCGGCTATCTTTATTTTGAACAAGTGCAAAAGATGCGCAAAAT 566
Db 421 TGTGACGGGACAAAGCGGCTATCTTTATTTTGAACAAGTGCAAAAGATGCGCAAAAT 480
QY 567 GTGCGACGCGCTTGGAGGAGCGTTTGAAGATTCTTGTACCGAGGATAGGTGAGT 626
Db 481 GTGCGACGCGCTTGGAGGAGCGTTTGAAGATTCTTGTACCGAGGATAGGTGAGT 540
QY 627 CATTTGATTGACAGACAGTCAATCTTACCGAAAGCCCAAGCTAGCTCATCTTGC 686
Db 541 CATTTGATTGACAGACAGTCAATCTTACCGAAAGCCCAAGCTAGCTCATCTTGC 600
QY 687 TGTGTA 692
Db 601 TGTGTA 606

RESULT 3
US-09-030-607-206/c
; Sequence 206, Application US/09030507
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 206:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 487 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

US-09-030-607-206

Query Match 37.4%; Score 439.6; DB 3; Length 487;
Best Local Similarity 96.0%; Pred. No. 2.3e-101;
Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
QY 536 TTTTGAACAAGTGCACAAAGATGCCACAAATGTGGCAGCGCTTTCAGGAAGCGGTTCG 595
Db 487 TTTTGAACAAGTGCACAAAGATGCCACAAATGTGGCAGCGCTTTCAGGAAGCGGTTCG 428
QY 596 AAGAGTTCTTCTCCGAGGATAGGTGAGATCAATTTGATTGAGATGTTGATGCTTC 655
Db 427 AAGAGTTCTTCTCCGAGGATAGGTGAGATCAATTTGATTGAGATGTTGATGCTTC 368
QY 656 TCACGGAAGCCCAAGCGCTAGC-TCACTCTTCTGCTTTCATTGTTAGATGTTGATGCTTC 714
Db 367 TCACGGAAGCCCAAGCGCTAGC-TCACTCTTCTGCTTTCATTGTTAGATGTTGATGCTTC 308
QY 715 TAACCAACTCACACATATACACAAATCAACATCGGGATGGAGAAGAGAATTAGCGTTTG 774
Db 307 TNACCAACTCACACATATACACATATACACATCGGGATGGAGAAGAGAATTAGCGTTTG 248
QY 775 CAGCAGTGATCATCTACTACTATAAAATTAACATTAATGCTGCTGCTTCATTAGTTGGTGG 834
Db 247 CAGCAGTGATCATCTACTACTATAAAATTAACATTAATGCTGCTGCTTCATTAGTTGGTGG 188
QY 835 AGAAGGACACATCCACTCTTTGGAGGAATATATTTACTCAATATGCGACCTTACATTTA 894
Db 187 AGAAGGACACATCCACTCTTTGGAGGAATATATNTACTCAATATGCGACCTTACATTTA 128
QY 895 TAAATTTGACAGTTGTCTAATAACGTTTCTTTTAAATTAATATGTAAGTTGACAGCTA 954
Db 127 TAAATTTGACAGTTGTCTAATAACGTTTCTTTTAAATTAATATGTAAGTTGACAGCTA 68
QY 955 ATAATGAAATGACCAAGCTTTAATTAATATAAATAAGAACTTGACTA 1006
Db 67 ATAATGAAATGACCAAGCTTTAATTAATATAAATAAGAACTTGACTA 16

RESULT 4

US-09-439-313-206/c
; Sequence 206, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqui
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(487)
; OTHER INFORMATION: n = A,T,C or G
US-09-439-313-206

Query Match

37.4%; Score 439.6; DB 4; Length 487;


```

RESULT 5
US-09-352-616A-206/c
; Sequence 206, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Xuqui
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF PROSTATE CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(487)
; OTHER INFORMATION: n = A,T,C or G
US-09-352-616A-206

Query Match 37.4%; Score 439.6; DB 4; Length 487;
Best Local Similarity 96.0%; Pred. No. 2.3e-101;
Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1

Qy 536 TTTTGAACAAGTGCAGGATGCCACAAATGTGGCAGCGCTTTGAGAGCGGTTCG 595
Db 487 TTTTGAACAAGTGCAGGCAANAGGCCACAAATGTGGCAGCGCTTTGAGAGCGGTTCG 428
Qy 596 AAGAGGTTCTTGCTACCGAGATAGTGCAGATCATTTGATTTCAGACACACACAGTCAATCT 655

```


QY 896 AATGCTACAGTGTCTATTAACGTTCTTTAAATTAATATGTAAGTTGAGAGCTAA 955
Db 123 AATTTGT-ACAGTTGTCTAA-ACGTTTCTTTAAATTAATATGTAAGTTGAGAGCTAA 66
QY 956 TAAATGAATGACCAAGACTTTTAAATATATAAATAAAGAACTTGACTA 1006
Db 65 TAAATGAATGACCAAGACTTTTAAATATATAAATAAAGAACTTGACTA 16

RESULT 11

US-09-232-149A-106/c
; Sequence 106, Application US/09232149A
; Patent No. 6465611

GENERAL INFORMATION:

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE
; FILE REFERENCE: 210121.42706
; CURRENT APPLICATION NUMBER: US/09/232,149A
; CURRENT FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 106
; LENGTH: 473
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-232-149A-106

Query Match 26.1%; Score 306.2; DB 4; Length 473;
Best Local Similarity 93.4%; Pred. No. 7.1e-68;
Matches 440; Conservative 0; Mismatches 18; Indels 13; Gaps 11;

QY 536 TTTTGAACAAGTGCACCAAGATGCCCAATGTGGCAGCAGCTTTGAGGAAGCGTTGC 595
Db 473 TTTTGAACAAGTGCACCAAGATGCCCAATGTGGCAGCAGCTTTGAGGAAGCGTTGC 417
QY 596 AAGAGTCTTGTACCGAGGATAGGTGCAGATCATTTGATTCAGACAGACAGTCAATCT 655
Db 416 AAGAGTCTTGTACCGAGGATAGGTGCAGATCATTTGATTCAGACAGACAGTCAATCT 357
QY 656 TCACCGAAGCCCAAGCTAGCTCATCTTGTCTGTGTTGATTTAGATTTGATGCAATCT 715
Db 356 TCACCGAAGCCCAAGCTAGCTCATCTTGTCTGTGTTGATTTAGATTTGATGCAATCT 299
QY 716 AACCAACTCACACATATACACAAATCAACATGGGATGGAGAGAAATTAGCGTTTCG 775
Db 298 -ACCAACTCACACATATACACATATC-ACATGGGATGGAGAGAAATTAGCGTTTCG 241
QY 776 AGCAGTGTATCATCTACTAATAAATTAACATAATGCTGCTTCATTAGTTGGTGGGA 835
Db 240 AGCAGTGTAT-ATCTACTAATAAATTAACATAATGCTGCTTCATTAGTTGGTGGGA 183
QY 836 GAAGGACACATCCCTCTGGGGATATATTTACTCAATATGCGACCTTACATTTAT 895
Db 182 GAAGGACACATCCCTCTGGGGATATATA-TTACTCAATATGCGACCTTACATTTAT 124
QY 896 AATTTGAACAGTGTCTAATAACGTTCTTTAAATTAATATGTAAGTTGAGAGCTAA 955
Db 123 AATTTGT-ACAGTTGTCTAA-ACGTTTCTTTAAATTAATATGTAAGTTGAGAGCTAA 66
QY 956 TAAATGAATGACCAAGACTTTTAAATATATAAATAAAGAACTTGACTA 1006
Db 65 TAAATGAATGACCAAGACTTTTAAATATATAAATAAAGAACTTGACTA 16

RESULT 12

US-09-016-434-1083
; Sequence 1083, Application US/09016434
; Patent No. 6500938

GENERAL INFORMATION:

; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Sellhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1083:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 624 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1174148
US-09-016-434-1083

Query Match 15.5%; Score 181.8; DB 4; Length 624;
Best Local Similarity 60.4%; Pred. No. 1.4e-36;
Matches 300; Conservative 0; Mismatches 197; Indels 0; Gaps 0;

QY 109 TTAAAGTAATCTCCTTGGAGATGGTGGAGAGTTTCACCTTATGAACAGATATG 168
Db 26 TGAAGGTATCATCTCTGGAGATTTCTGGAGTCGGGAAGACATCACTCATGAACAGATG 85
QY 169 TAACATAAAGTTTGATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTAAATAAG 228
Db 86 TGAATAAGAAATTCAGCAATCAGTACAAAGCCCAATAGGAGCTGACTTTTGACCAAG 145
QY 229 ATTTGGAAGTGGATGAGCATTTGTTACCATGAGATTTGGGACAGCGAGGTGAGGAGC 288
Db 146 AGGTGATGGTGGATGACAGGCTGCTGCAATGAGATATGGGACACAGCAGGACAGGAAC 205
QY 289 GATTCGGAAGCCTGAGGACACCATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTA 348
Db 206 GTTTCAGTCTCTCGGTGTGGCCCTTCTACAGAGGTGAGAGTCTGCTGCTGCTGTTT 265
QY 349 GTGTCGATGATTCACAAAGCTTCAGAACTTAAGTAACCTGGAAGAAAGAAATTCATATAT 408
Db 266 ATGTGACTGCCCCCAACACATTCAAAACCCCTAGATAGCTGGAGAGATGAGTTCTCTGCTC 325
QY 409 ATGCAGATGTGAAGAGCCTGAGAGCTTTCTCTTTTGTGATCTTGGGTAAACAAGTTGACA 468
Db 326 AGGCCAGTCCCCGAGATCCCTGAAACATTTCCCATTTGTTGTTGGGAAAACAAGTTGACC 385
QY 469 TAAGCGAACCGCAGGTGTCTACAGAAAGAACCCCAAGCTTGTGTCAGGACCAACGGCGACT 528

Best Local Similarity 53.5%; Pred. No. 2.7e-11;
Matches 170; Conservative 0; Mismatches 148; Indels 0; Gaps 0;

QY 106 TTTTAAAGTAATCTCTCGAGATGCTGGAGTGGAGAGTTCATTATGACAGAT 165
Db 82 TATTCAGATTACTCTGATTTGGCACTCAGGGGTGGAAAGTCTTGCTTCTTTAGGT 141
QY 166 ATGTAACATAAATAGTTTATGATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTAAATA 225
Db 142 TTGCAGATGATACATATACAGAAAGTACATCAGACAAATGTTGTTGATTTCAAAATAA 201
QY 226 AAGATTGGAGTGGATGCAATTTTGTATCCATGCGAGATTGGGACACGGCAGGTGAGG 285
Db 202 GAATCATAGATTAGACGGGAAACAAATCAAGCTTCAATATATGGGACACAGCGCCAGG 261
QY 286 AGCATTCGGAAGCTCGAGGACACCAATTTACAGAGTCTGACTGCTGCTGCTTACTT 345
Db 262 AAGATTTCGAACATCACCTCCAGTTATTACAGAGAGCCCATGGATCATATGTTGTGT 321
QY 346 TTAGTGTGATGATTCACAAAGCTTCCAGAACTTAACTGAAAGAAAGAAATTCATAT 405
Db 322 ATGATGTGACAGATCAGAGTCTTCAATAATGTTAAACAGTGGCTGCGGAAATAGATC 381
QY 406 ATTATGAGATGTGAAAG 423
Db 382 GTTATGCGAGTGAATG 399

RESULT 15

US-09-016-434-1060
; Sequence 1060, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1060:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 944 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK

CLONE: g1055280
US-09-016-434-1060

Query Match 5.9%; Score 69.6; DB 4; Length 944;
Best Local Similarity 53.4%; Pred. No. 2.4e-08;
Matches 175; Conservative 0; Mismatches 144; Indels 9; Gaps 1;

QY 257 CATGCAGATTTGGGACACGGCAGGTTCAGAGCGATTCCGAGCCTTGAGGACACCACTTTTA 316
Db 452 CTTGAGTTTATGGGACACAGCGGGCAGGAGGTTTCGTAGCTTAACGACACAGCGCTTTT 511
QY 317 CAGAGGTTCTGACTGCTGCTGCTTACTTTAGTGTGATGATTCAAAAGCTTCCAGAA 376
Db 512 CAGAGATGCTATGGGTTTCTTCTTACTTTGATCTGACAAATGAGCAAGTTTCTCTCA 571
QY 377 CTTAAGTAACTGGAAGAAGAATTCATATATATGACAGATGTGAAGAGCCTGAGAGCTT 436
Db 572 TGTCAAGAACTGGATATAGCCAGCTACAGATGCTATGTCATATTTGTGAAAACCCAGA 625
QY 437 TCCTTTTGTGATTTCTGGGTAAACAAGATTGACATAAGCGAAGCGCAGGTGTCTACAGAGA 496
Db 626 ---TATAGTGTGTTGGAAACAGAGTGTCTGGAGGACCGAGAGTAGTGAAGAGGA 682
QY 497 AGCCCAAGCTTGGTGCAGGGCAACAGCGGCACTATCTTTTATTTTGAACAAGTGCAGAAA 556
Db 683 GGAAGCCATAGCACTTCGCAGAGAAATATGGAATCCCTTACTTTGAAACTAGTGTGCTG 742
QY 557 TGCCCAAAATGTGGCAGCAGCGCTTTGAG 584
Db 743 TGGGACAAACATAGCCCAAGCAATTGAG 770

Search completed: February 10, 2004, 04:46:01
Job time : 97 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2004, 02:32:31 ; Search time 489 Seconds

(without alignments)
8851.265 Million cell updates/sec

Title: US-09-988-974-6

Perfect score: 1175

Sequence: 1 CACTGTGATGAACACTTTT.....GTAAAAA.....1175

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2449703 seqs, 1841816367 residues

Total number of hits satisfying chosen parameters: 4899406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata1/pubpna/PCTU5_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata1/pubpna/US09_PUB.seq.*
- 14: /cgn2_6/ptodata1/pubpna/US10A_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata1/pubpna/US10_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata1/pubpna/US10_NEW_PUB.seq.*
- 17: /cgn2_6/ptodata1/pubpna/US60_NEW_PUB.seq.*
- 18: /cgn2_6/ptodata1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1175	100.0	1175	9	US-09-988-974-6
2	1160.2	98.7	1372	13	Sequence 6, Appli
3	1100	93.6	1969	9	Sequence 488, App
4	1100	93.6	6058	11	Sequence 10308, A
5	632	53.8	632	9	Sequence 5493, A
6	604.4	51.4	606	12	Sequence 26942, A
7	439.6	37.4	487	9	Sequence 1082, Ap
8	439.6	37.4	487	9	Sequence 206, App
9	439.6	37.4	487	9	Sequence 206, App
10	439.6	37.4	487	9	Sequence 206, App
11	439.6	37.4	487	9	Sequence 206, App
12	439.6	37.4	487	10	Sequence 206, App
13	439.6	37.4	487	10	Sequence 206, App
14	439.6	37.4	487	10	Sequence 206, App
15	439.6	37.4	487	13	Sequence 206, App

c 16	439.6	37.4	487	13	US-10-294-025-206	Sequence 206, App
c 17	439.6	37.4	487	14	US-10-012-896-206	Sequence 206, App
c 18	439.6	37.4	487	15	US-10-010-940-206	Sequence 206, App
c 19	394.4	33.6	397	12	US-10-242-535A-42720	Sequence 42720, A
c 20	356	30.3	393	13	US-10-101-510-138	Sequence 138, App
c 21	333.2	28.4	648	13	US-10-029-386-24799	Sequence 24799, A
c 22	333.2	28.4	1114	10	US-09-764-868-508	Sequence 508, App
c 23	322.4	27.4	1111	10	US-09-764-868-89	Sequence 89, Appli
c 24	306.2	26.1	473	9	US-09-759-143-106	Sequence 106, App
c 25	306.2	26.1	473	9	US-09-780-869-106	Sequence 106, App
c 26	306.2	26.1	473	9	US-09-030-606-106	Sequence 106, App
c 27	306.2	26.1	473	9	US-09-822-827-106	Sequence 106, App
c 28	306.2	26.1	473	9	US-09-115-453-106	Sequence 106, App
c 29	306.2	26.1	473	10	US-09-232-880-106	Sequence 106, App
c 30	306.2	26.1	473	10	US-09-895-793-106	Sequence 106, App
c 31	306.2	26.1	473	10	US-09-895-814-106	Sequence 106, App
c 32	306.2	26.1	473	13	US-10-144-678A-106	Sequence 106, App
c 33	306.2	26.1	473	13	US-10-294-025-106	Sequence 106, App
c 34	306.2	26.1	473	14	US-10-012-896-106	Sequence 106, App
c 35	306.2	26.1	473	15	US-10-010-940-106	Sequence 106, App
c 36	188.4	16.0	210	12	US-10-242-535A-93	Sequence 93, Appli
c 37	185.8	15.8	462	9	US-09-864-761-2267	Sequence 2267, Ap
c 38	183.4	15.6	2178	9	US-09-925-302-186	Sequence 186, App
c 39	183.4	15.6	2302	12	US-10-264-049-724	Sequence 724, App
c 40	181.8	15.5	624	12	US-10-305-720-1083	Sequence 1083, Ap
c 41	141.2	12.0	627	12	US-10-369-493-25550	Sequence 25550, A
c 42	137.4	11.7	698	9	US-09-828-310-8	Sequence 8, Appli
c 43	133.4	11.4	223	9	US-09-864-761-19005	Sequence 19005, A
c 44	128.4	10.9	941	9	US-09-770-445-358	Sequence 358, App
c 45	128.2	10.9	612	10	US-09-938-842A-1196	Sequence 1196, Ap

ALIGNMENTS

RESULT 1

US-09-988-974-6
; Sequence 6, Application US/09988974
; Patent No. US20020090712A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Goli, Surya K.
; Bandman, Olga
; TITLE OF INVENTION: NOVEL RAB PROTEINS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: fastseq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/988,974
; FILING DATE: 19-No. US20020090712A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/215,887
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0183 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 6:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1175 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-988-974-6

Query Match      100.0%; Score 1175; DB 9; Length 1175;
Best Local Similarity 100.0%; Pred. No. 5.5e-253;
Matches 1175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACTGTGATGAACACTTTCCCGTGTGCTTGGATGCACTCTTCAACAACCCCTAGGAG 60
DB 1 CACTGTGATGAACACTTTCCCGTGTGCTTGGATGCACTCTTCAACAACCCCTAGGAG 60
QY 61 GGTCTCTGAAGCTTTTGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC 120
DB 61 GGTCTCTGAAGCTTTTGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC 120
QY 121 TCCTTGGAGATGGTGGAGTGGGAAGATCTCTATGACAGATATGTAATAAAGT 180
DB 121 TCCTTGGAGATGGTGGAGTGGGAAGATCTCTATGACAGATATGTAATAAAGT 180
QY 181 TTGATACCCAGCTCTCCATCAATAGGTGTGGAATTTTAAATAAGATTTGGAAGTGG 240
DB 181 TTGATACCCAGCTCTCCATCAATAGGTGTGGAATTTTAAATAAGATTTGGAAGTGG 240
QY 241 ATGGACATTTTGTACCATGAGATTTGGGACACGGCAGGTGAGGAGGATTCGGAAGCC 300
DB 241 ATGGACATTTTGTACCATGAGATTTGGGACACGGCAGGTGAGGAGGATTCGGAAGCC 300
QY 301 TGAGGACACATTTTACAGAGTTCTGACTGTGCTGCTTACTTTTATGTCGATGATT 360
DB 301 TGAGGACACATTTTACAGAGTTCTGACTGTGCTGCTTACTTTTATGTCGATGATT 360
QY 361 CACAAAGCTTCCAGAACCTTAAGTAACTGGAAGAAAGATTCATATATTATGCAATGTA 420
DB 361 CACAAAGCTTCCAGAACCTTAAGTAACTGGAAGAAAGATTCATATATTATGCAATGTA 420
QY 421 AAGAGCTGAGAGCTTTCCCTTTTGTGATTTCTGGGTAAACAAGTTGACATGCAAGCCG 480
DB 421 AAGAGCTGAGAGCTTTCCCTTTTGTGATTTCTGGGTAAACAAGTTGACATGCAAGCCG 480
QY 481 AGGTGCTACAGAGAGCCCAAGCTTGGTGCAGGACAAACGGGACTATCTTATTTTG 540
DB 481 AGGTGCTACAGAGAGCCCAAGCTTGGTGCAGGACAAACGGGACTATCTTATTTTG 540
QY 541 AAACAAGTCCAAAGATGCCAATAATGTGGCAGCAGCCTTTGAGGAAGCGGTTCGAAGAG 600
DB 541 AAACAAGTCCAAAGATGCCAATAATGTGGCAGCAGCCTTTGAGGAAGCGGTTCGAAGAG 600
QY 601 TTCTTGCTACCCAGGATAGTGCATCTATTCATTCAGACACACAGTCAATCTTACC 660
DB 601 TTCTTGCTACCCAGGATAGTGCATCTATTCATTCAGACACACAGTCAATCTTACC 660
QY 661 GAAAGCCCAAGCTAGCTCATCTTGTGTTGATTGTTAGATTGTTGATGCTTCAACCA 720
DB 661 GAAAGCCCAAGCTAGCTCATCTTGTGTTGATTGTTAGATTGTTGATGCTTCAACCA 720
QY 721 ACTCACATATACACAAATCAACATGGGGATGGAGAGAAATAGCGTTGCGAGAG 780
DB 721 ACTCACATATACACAAATCAACATGGGGATGGAGAGAAATAGCGTTGCGAGAG 780
QY 781 TGTATCATCTACTAATAAATTAATTAATGCTGCTTCAATAGTTGGTGGGAGAGG 840
DB 781 TGTATCATCTACTAATAAATTAATTAATGCTGCTTCAATAGTTGGTGGGAGAGG 840
QY 841 GACACATCCACTTTGGAGAAATATATTTACTCAATATGCAATGCAACCTTCAATTAAT 900
DB 841 GACACATCCACTTTGGAGAAATATATTTACTCAATATGCAATGCAACCTTCAATTAAT 900

```

```

DB 841 GACACATCCACTTTGGAGAAATATATTTACTCAATATGCAACCTTCAATTAATTAAT 900
QY 901 GTAAACAGTTGTCTAATACGTTTCTTTTAATTTAAATATGTAAGTTGCAGAGCTAATAAT 960
DB 901 GTAAACAGTTGTCTAATACGTTTCTTTTAATTTAAATATGTAAGTTGCAGAGCTAATAAT 960
QY 961 GAAATGACCAAGACTTTTAATTAATAAATAAAGAACTTGCATCTTCTAGAAGTTATA 1020
DB 961 GAAATGACCAAGACTTTTAATTAATAAATAAAGAACTTGCATCTTCTAGAAGTTATA 1020
QY 1021 CTTGGATTTTTCCTGGGAAATGGAGAACTACTTTTATATGTGTATGTTTATGCAA 1080
DB 1021 CTTGGATTTTTCCTGGGAAATGGAGAACTACTTTTATATGTGTATGTTTATGCAA 1080
QY 1081 TTAGCATTTCTATTCTGTTTTCAGGAAATACCTTCTTAAGCAATAATGTTAGATATAA 1140
DB 1081 TTAGCATTTCTATTCTGTTTTCAGGAAATACCTTCTTAAGCAATAATGTTAGATATAA 1140
QY 1141 AGATTAAATCTAATGTAATAAATAAATAAATAAATAA 1175
DB 1141 AGATTAAATCTAATGTAATAAATAAATAAATAAATAA 1175

RESULT 2
US-10-101-510-488
; Sequence 488, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 488
; LENGTH: 1372
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: modified_base
; LOCATION: (1362)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1364)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1366)
; OTHER INFORMATION: a, t, c, g, other or unknown
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1368)
; OTHER INFORMATION: a, t, c, g, other or unknown
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-488

Query Match      98.7%; Score 1160.2; DB 13; Length 1372;
Best Local Similarity 99.2%; Pred. No. 1.2e-249;
Matches 1165; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 CACTGTGATGAACACTTTTCCCGTGTGCTTGGATGCACTCTTCAACAACCCCTAGGAG 60
DB 190 CACTGTGATGAACACTTTTCCCGTGTGCTTGGATGCACTCTTCAACAACCCCTAGGAG 249
QY 61 GGTCTCTGAAGCTTTTGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC 120
DB 250 GGTCTCTGAAGCTTTTGAGATTAAACAATGGCAGGAAATCATCACTTTTAAAGTAATTC 309
QY 121 TCCTTGGAGATGGTGGAGTGGGAAGATTCATTTATGACAGATATGTAATAAAGT 180

```


310 TCCTGGAGATGGTGGAGTTCGGAGAGTTCACCTATGACAGATATGTAACATTAAGT 369
181 TTGATACCCAGCTCTTCCATCAATAGTGTGGAAATTTTAAATAAAGATTTGGAAGTGG 240
370 TTGATACCCAGCTCTTCCATCAATAGTGTGGAAATTTTAAATAAAGATTTGGAAGTGG 429
241 ATGGACATTTTGTGTACCATGACAGATTTGGGACACGGCAGGTTCAGGAGCGATTCGGAAGCC 300
430 ATGGACATTTTGTGTACCATGACAGATTTGGGACACGGCAGGTTCAGGAGCGATTCGGAAGCC 489
301 TGAGGACACCATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGATT 360
490 TGAGGACACCATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGATT 549
361 CACAAGCTTCCAGAACTTAACTAGTAACTGGAAGAAAGAAATCATATATATGCGAGATGTA 420
550 CACAAGCTTCCAGAACTTAACTAGTAACTGGAAGAAAGAAATCATATATATGCGAGATGTA 609
421 AAGAGCTGAGAGCTTCCCTTTTGTGATTTGGGTAAACAAGATTGACATTAAGCGAACGGC 480
610 AAGAGCTGAGAGCTTCCCTTTTGTGATTTGGGTAAACAAGATTGACATTAAGCGAACGGC 669
481 AGGTGCTTACAGAAAGCCAGCTTGTGCGAGGACACGGGCTACTATCTTATTTTG 540
670 AGGTGCTTACAGAAAGCCAGCTTGTGCGAGGACACGGGCTACTATCTTATTTTG 729
541 AAACAAGTGAAGAGTGCACAAATGTGCAGCAGCCTTTGAGGAAGCGTTTCGAAGAG 600
730 AAACAAGTGAAGAGTGCACAAATGTGCAGCAGCCTTTGAGGAAGCGTTTCGAAGAG 789
601 TTCTTGCTACCGAGGATAGTCAGATCAATTTGATTCAGACAGACAGCAATCTTCACC 660
790 TTCTTGCTACCGAGGATAGTCAGATCAATTTGATTCAGACAGACAGCAATCTTCACC 849
661 GAAAGCCCAAGCTAGCTCATCTGCTGTTGATTTGTAGATTCTGATGATCTCAACCA 720
850 GAAAGCCCAAGCTAGCTCATCTGCTGTTGATTTGTAGATTCTGATGATCTCAACCA 909
721 ACTCACATATACACAAATCAACATGGGGATGGAGAGAGAAATTAGCGTTTCGACGAG 780
910 ACTCACATATACACAAATCAACATGGGGATGGAGAGAGAAATTAGCGTTTCGACGAG 969
781 TGTATCATCTACTAATAAATTAACCTAATGTGCTGCTCATTAGTTGGTGGGAGAGG 840
970 TGTATCATCTACTAATAAATTAACCTAATGTGCTGCTCATTAGTTGGTGGGAGAGG 1029
841 GACATATCCACTCTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTATAAAT 900
1030 GACATATCCACTCTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTATAAAT 1089
901 GTAACAGTTGCTAATAACGTTTCTTTAATTAATGATGATTAAGTTCGACGAGCTATAAAT 960
1090 GTAACAGTTGCTAATAACGTTTCTTTAATTAATGATGATTAAGTTCGACGAGCTATAAAT 1149
961 GAAATGACCAAGACTTTAATTAATAAATAAAGAACTTGACTATTTCTAGAAGTTATA 1020
1150 GAAATGACCAAGACTTTAATTAATAAATAAAGAACTTGACTATTTCTAGAAGTTATA 1209
1021 CTGGAATTTTCTGGGAAATGAGAACTATTTTATATGATGATGTTTTATGCAA 1080
1210 CTGGAATTTTCTGGGAAATGAGAACTATTTTATATGATGATGTTTTATGCAA 1269
1081 TTAGCATTTGATTTCTGGTTTCAGGAAATATCTTCTTAAGCAATTAATGTTAGATTTAA 1140
1270 TTAGCATTTGATTTCTGGTTTCAGGAAATATCTTCTTAAGCAATTAATGTTAGATTTAA 1329
1141 AGATTAATAATCTAATGTAATAAAAAAAAAAAAAA 1174
1330 AGATTAATAATCTAATGTAATTTGCAATGCAAAAANA 1363

RESULT 3

US-09-864-761-10308
; Sequence 10308, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 10308
; LENGTH: 1969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC003037.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
US-09-864-761-10308
Query Match 93.6%; Score 1100; DB 9; Length 1969;
Best Local Similarity 100.0%; Pred. No. 4.4e-236;
Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 AGGGTCTTGAAGCTTTTGAGATTAAACAATGGCAGGAATCATCACTTTTAAAGTAAT 118
DB 311 AGGGTCTTGAAGCTTTTGAGATTAAACAATGGCAGGAATCATCACTTTTAAAGTAAT 370
QY 119 TCTCCTGGAGATGGTGGAGTTCGGAGAGTTCACCTTATGAAACAGATATGTAACATAA 178

```

Db 371 TCCTCTGGAGATGGTGGAGAGTTCATTTATGACAGATATGTAATTAATAA 430
Qy 179 GTTTGATACCCAGCTCTCCATACAAATAGTGTGGAAATTTTAAATAAGATTTGAAGT 238
Db 431 GTTTGATACCCAGCTCTCCATACAAATAGTGTGGAAATTTTAAATAAGATTTGAAGT 490
Qy 239 GGATGGACATTTTGTACCATGACAGATTTGGGACAGGAGGTCAGGAGGATTCGGAAG 298
Db 491 GGATGGACATTTTGTACCATGACAGATTTGGGACAGGAGGTCAGGAGGATTCGGAAG 550
Qy 299 CTTGAGGACACCATTTTACAGAGGTTCTGATGCTGCTCTCTCTCTCTCTCTCTCTCT 358
Db 551 CTTGAGGACACCATTTTACAGAGGTTCTGATGCTGCTCTCTCTCTCTCTCTCTCTCT 610
Qy 359 TTCAAAAGCTTCCAGAACTTAACTAGTAACTGGAAGAGATTCATATATATGCAATGT 418
Db 611 TTCAAAAGCTTCCAGAACTTAACTAGTAACTGGAAGAGATTCATATATATGCAATGT 670
Qy 419 GAAAGAGCTTGAGAGCTTTCTTTTGTGATTTCTGGGTAAACAGATTCACATAGCGAAG 478
Db 671 GAAAGAGCTTGAGAGCTTTCTTTTGTGATTTCTGGGTAAACAGATTCACATAGCGAAG 730
Qy 479 GCAGGTGTCTACAGAAAGCCAGCTTGGTCAGGACACCGGCACTATCTCTTATTT 538
Db 731 GCAGGTGTCTACAGAAAGCCAGCTTGGTCAGGACACCGGCACTATCTCTTATTT 790
Qy 539 TGAACAAGTGCAGAAAGATGCCAATGTGGCAGAGCTTTGAGGAAGCGGTTCCGAG 598
Db 791 TGAACAAGTGCAGAAAGATGCCAATGTGGCAGAGCTTTGAGGAAGCGGTTCCGAG 850
Qy 599 AGTTCTGCTACGAGGATAGGTGAGATCAATTTGATTCAGACAGACAGTCAATCTCA 658
Db 851 AGTTCTGCTACGAGGATAGGTGAGATCAATTTGATTCAGACAGACAGTCAATCTCA 910
Qy 659 CCGAAGCCCAAGCTAGCTCATCTTCTGCTGATTTGATTTGATTTGATTTGATTTGAT 718
Db 911 CCGAAGCCCAAGCTAGCTCATCTTCTGCTGATTTGATTTGATTTGATTTGATTTGAT 970
Qy 719 CAATCTCACATATACACAAATCAATGGGATGGAGAGAGAAATAGCGTTTGCAGC 778
Db 971 CAATCTCACATATACACAAATCAATGGGATGGAGAGAGAAATAGCGTTTGCAGC 1030
Qy 779 AGGTATCTACTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 838
Db 1031 AGGTATCTACTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1090
Qy 839 GGGACACATCCACTCTTGGAGGATATATTTACTCAATAATGGCACTTACATTTATATA 898
Db 1091 GGGACACATCCACTCTTGGAGGATATATTTACTCAATAATGGCACTTACATTTATATA 1150
Qy 899 TTGTAACAGTTGTCTAATAACGTTTCTTTAATTAATATGTAAGTTGAGAGCTTAATA 958
Db 1151 TTGTAACAGTTGTCTAATAACGTTTCTTTAATTAATATGTAAGTTGAGAGCTTAATA 1210
Qy 959 ATGAAATGACCAAGACTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1018
Db 1211 ATGAAATGACCAAGACTTTAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 1270
Qy 1019 TACTTGGAATTTTCTCGGAAATAGGAACTPACTTTTTATATGATGATGTTTTATGTC 1078
Db 1271 TACTTGGAATTTTCTCGGAAATAGGAACTPACTTTTTATATGATGATGTTTTATGTC 1330
Qy 1079 AATTAGCATTTGATTTCTGTTGAGGAAATCTTTTCCTAAGCAATATGATGATTTAT 1138
Db 1331 AATTAGCATTTGATTTCTGTTGAGGAAATCTTTTCCTAAGCAATATGATGATTTAT 1390
Qy 1139 AAAGATTAATAATCTAATGTA 1158
Db 1391 AAAGATTAATAATCTAATGTA 1410

```

RESULT 4

```

US-09-764-891-5493
; Sequence 5493, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5493
; LENGTH: 6058
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-5493

```

Query Match 93.6%; Score 1100; DB 11; Length 6058;

Best Local Similarity 100.0%; Pred. No. 7.4e-236;

Matches 1100; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 59 AGGTTCTTGAAGCTTTTGAAGTTAAACATGGCGAGGAAATCATCTTTTAAAGTAAT 118
Db 4905 AGGTTCTTGAAGCTTTTGAAGTTAAACATGGCGAGGAAATCATCTTTTAAAGTAAT 4964
Qy 119 TCTCTTGGAGATGGTGGAGTTGGAGAGTTTCACTTATGAACAGATATGTAATAATAA 178
Db 4965 TCTCTTGGAGATGGTGGAGTTGGAGAGTTTCACTTATGAACAGATATGTAATAATAA 5024
Qy 179 GTTTGATATCCAGCTCTTCCATACATAGGTGGAAATTTTAAATAAAGATTTGGAAGT 238
Db 5025 GTTTGATATCCAGCTCTTCCATACATAGGTGGAAATTTTAAATAAAGATTTGGAAGT 5084
Qy 239 GGATGGACATTTTCTTACCATGAGATTTGGGACAGCGAGTCAGAGCGATTCGGAAG 298
Db 5085 GGATGGACATTTTCTTACCATGAGATTTGGGACAGCGAGTCAGAGCGATTCGGAAG 5144
Qy 299 CTTGAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTGTAGTGTGATGA 358
Db 5145 CTTGAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTGTAGTGTGATGA 5204
Qy 359 TTCAAAAGCTTCCAGAACTTAACTAGTAACTGGAAGAGAAATTCATATATATGCAATGT 418
Db 5205 TTCAAAAGCTTCCAGAACTTAACTAGTAACTGGAAGAGAAATTCATATATATGCAATGT 5264
Qy 419 GAAAGAGCTTGAGAGCTTTCTTTTGTGATTTCTGGGTAAACAGATTCACATAGCGAAG 478
Db 5265 GAAAGAGCTTGAGAGCTTTCTTTTGTGATTTCTGGGTAAACAGATTCACATAGCGAAG 5324
Qy 479 GCAGGTGTCTACAGAAAGCCCAAGCTTGGTGCAGGGAACAACGGGACTATCTCTTATTT 538
Db 5325 GCAGGTGTCTACAGAAAGCCCAAGCTTGGTGCAGGGAACAACGGGACTATCTCTTATTT 5384
Qy 539 TGAACAAGTGCAGAAAGATGCCAATATGTCAGACAGGCTTTGAGGAAGCGGTTCCGAG 598
Db 5385 TGAACAAGTGCAGAAAGATGCCAATATGTCAGACAGGCTTTTGGAGAGCGGTTCCGAG 5444
Qy 599 AGTTCTTGTCTACCGAGGATAGGTGAGATCAATTTGATTCAGACAGACACAGTCAATCTTCA 658
Db 5445 AGTTCTTGTCTACCGAGGATAGGTGAGATCAATTTGATTCAGACAGACACAGTCAATCTTCA 5504
Qy 659 CCGAAGCCCAAGCTAGCTCATCTTCTGCTGATTTGATTTGATTTGATTTGATTTGAT 718
Db 5505 CCGAAGCCCAAGCTAGCTCATCTTCTGCTGATTTGATTTGATTTGATTTGATTTGAT 5564
Qy 719 CAATCTCACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCGTTTTCGAGC 778
Db 5565 CAATCTCACATATACACAAATCAACATGGGATGGAGAGAGAAATAGCGTTTTCGAGC 5624
Qy 779 AGTGTATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 838
Db 5625 AGTGTATCTACTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5684

```

839 GGACACATCCACTCTTGGAGGAATATATTACTCAATATGACCTTACATTTATAA 898
 Db GGACACATCCACTCTTGGAGGAATATATTACTCAATATGACCTTACATTTATAA 5744
 899 TTGTAACAGTGTCTAATAAGCTTCTTTAATTAAATATGTAAGTTGACAGCTAATAA 958
 Db TTGTAACAGTGTCTAATAAGCTTCTTTAATTAAATATGTAAGTTGACAGCTAATAA 5804
 959 ATGAATGACCAAGCACTTAAATATATAAATAAGAACTGCACTATTCTAGAGTTA 1018
 Db ATGAATGACCAAGCACTTAAATATATAAATAAGAACTGCACTATTCTAGAGTTA 5864
 1019 TACTTGGATTTTCTCGGAAATGAGAACTACTTTTTATATGTAATGTTTATGTC 1078
 Db TACTTGGATTTTCTCGGAAATGAGAACTACTTTTTATATGTAATGTTTATGTC 5924
 1079 AATTAGCAATGTAATCTTCTGTTGAGGAAATACCTTCTTAAAGCAATAATGTTAGATATT 1138
 Db AATTAGCAATGTAATCTTCTGTTGAGGAAATACCTTCTTAAAGCAATAATGTTAGATATT 5984
 1139 AAAGATTAAATCTAATGTA 1158
 Db AAAGATTAAATCTAATGTA 6004

RESULT 5
 US-09-864-761-26942
 ; Sequence 26942, Application US/09864761
 ; Patent No. US20020048763A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng
 ; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 ; FILE REFERENCE: Reomica-X-1
 ; CURRENT APPLICATION NUMBER: US/09/864,761
 ; CURRENT FILING DATE: 2001-05-23
 ; PRIOR APPLICATION NUMBER: US 60/180,312
 ; PRIOR FILING DATE: 2000-02-04
 ; PRIOR APPLICATION NUMBER: US 60/207,456
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: US 09/632,366
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: GB 24263.6
 ; PRIOR FILING DATE: 2000-10-04
 ; PRIOR APPLICATION NUMBER: US 60/236,359
 ; PRIOR FILING DATE: 2000-09-27
 ; PRIOR APPLICATION NUMBER: PCT/US01/00666
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00667
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00664
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00669
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00665
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00668
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00663
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00662
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00661
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: PCT/US01/00670
 ; PRIOR FILING DATE: 2001-01-30
 ; PRIOR APPLICATION NUMBER: US 60/234,687
 ; PRIOR FILING DATE: 2000-09-21
 ; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: US 09/774,203
 ; PRIOR FILING DATE: 2001-01-29
 ; NUMBER OF SEQ ID NOS: 49117
 ; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
 ; SEQ ID NO 26942
 ; LENGTH: 632
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: MAP TO AC003037.1
 ; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
 ; OTHER INFORMATION: EXPRESSED IN HELI00, SIGNAL = 1.2
 ; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
 ; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.2
 ; OTHER INFORMATION: EST HUMAN HIT: BE735344.1, EVALU0 0.00e+00
 ; OTHER INFORMATION: NT HIT: g11421053, EVALU0 0.00e+00
 ; OTHER INFORMATION: SWISSPROT HIT: P51151, EVALU0 1.00e-108
 ; US-09-864-761-26942

Query Match 53.8%; Score 632; DB 9; Length 632;
 Best Local Similarity 100.0%; Pred. No. 1.6e-131;
 Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 61 GGTCTTGAAGCTTTTGAGATTAAACAATGCGCAGGAAATCATCATCTTTTAAAGTAATTC 120
 Db 1 GGTCTTGAAGCTTTTGAGATTAAACAATGCGCAGGAAATCATCATCTTTTAAAGTAATTC 60
 QY 121 TCGTTGGAGATGGTGGAGTTGCGGAAGAGTTCACATTATGAACAGATATGTAACATAAAGT 180
 Db 61 TCGTTGGAGATGGTGGAGTTGCGGAAGAGTTCACATTATGAACAGATATGTAACATAAAGT 120
 QY 181 TTGATACCCAGCTCTTCATACATAGTGTGGAAATTTTAAATAAAGATTTGGAAGTGG 240
 Db 121 TTGATACCCAGCTCTTCATACATAGTGTGGAAATTTTAAATAAAGATTTGGAAGTGG 180
 QY 241 ATGACACATTTTGTACCATGAGATTGGGACACGAGGTCAGGAGCGATTCCGAGGCC 300
 Db 181 ATGACACATTTTGTACCATGAGATTGGGACACGAGGTCAGGAGCGATTCCGAGGCC 240
 QY 301 TGAGGACACCATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGAT 360
 Db 241 TGAGGACACCATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTTAGTGTGATGAT 300
 QY 361 CACAAAGCTTCCAGAACTTAAAGTAACTGGAAGAAAGAAATTCATATATATGACATGTA 420
 Db 301 CACAAAGCTTCCAGAACTTAAAGTAACTGGAAGAAAGAAATTCATATATATGACATGTA 360
 QY 421 AAGAGCTGAGAGCTTTCTTTTGTGATTCTGGGTAAACAAGATTGACATGAAGCGAGCGC 480
 Db 361 AAGAGCTGAGAGCTTTCTTTTGTGATTCTGGGTAAACAAGATTGACATGAAGCGAGCGC 420
 QY 481 AGGTGTCTACAGAGAGAGCCCAAGCTTGGTGCAGGAGCAACGGGACTATCTTTATTG 540
 Db 421 AGGTGTCTACAGAGAGAGCCCAAGCTTGGTGCAGGAGCAACGGGACTATCTTTATTG 480
 QY 541 AAACAAGTGCAAAAGATGCCACAAATGTGGCAGCAGCTTTTGAGGAAGCGGTTCGAAGAG 600
 Db 481 AAACAAGTGCAAAAGATGCCACAAATGTGGCAGCAGCTTTTGAGGAAGCGGTTCGAAGAG 540
 QY 601 TTCTTGTACCGAGGATAGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCTTCAAC 660
 Db 541 TTCTTGTACCGAGGATAGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCTTCAAC 600
 QY 661 GAAAGCCCAAGCCCTAGCTCATCTTGTGTTGA 692
 Db 601 GAAAGCCCAAGCCCTAGCTCATCTTGTGTTGA 632

RESULT 6
 US-10-305-720-1082
 ; Sequence 1082, Application US/10305720

; Publication No. US20040010136A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Au-Young, Janice K.; Sellhamer, Jeffrey J.
 ; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
 ; FILE REFERENCE: PA-0002-1 CON
 ; CURRENT APPLICATION NUMBER: US/10/305,720
 ; CURRENT FILING DATE: 2002-11-26
 ; PRIOR APPLICATION NUMBER: 09/016,434
 ; PRIOR FILING DATE: 1998-01-30
 ; NUMBER OF SEQ ID NOS: 1490
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 1082
 ; LENGTH: 606
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: GenBank ID No. US20040010136A1 g1174146
 US-10-305-720-1082

Query Match 51.4%; Score 604.4; DB 12; Length 606;
 Best Local Similarity 99.8%; Pred. No. 2.3e-125;
 Matches 605; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	87	ATGCGAGGAAAATCATCATTCTTTTAAAGTAAATCTCTTGGAGATGCTGAGTTGGAAG	146
Db	1	ATGCGAGGAAAATCTTCCATCTTTTAAAGTAAATCTCTTGGAGATGCTGAGTTGGAAG	60
Qy	147	AGTTCACTTATGAACAGATATGTAACTAATTAAGTTTGATACCCAGCTCTTCCATACAATA	206
Db	61	AGTTCACTTATGAACAGATATGTAACTAATTAAGTTTGATACCCAGCTCTTCCATACAATA	120
Qy	207	GGTGTGGAATTTTAAATAAAGATTGGAAGTGAATGACATTTTGTTACCATGCAGATT	266
Db	121	GGTGTGGAATTTTAAATAAAGATTGGAAGTGAATGACATTTTGTTACCATGCAGATT	180
Qy	267	TGGGACACGCGCAGCTCAGGAGCGATTCCGAAGCCTGAGGACACCATTTTACAGAGTTCT	326
Db	181	TGGGACACGCGCAGCTCAGGAGCGATTCCGAAGCCTGAGGACACCATTTTACAGAGTTCT	240
Qy	327	GACTGTGCTGCTTACTTTTAGTGTGATGATTCACAAAGCTTCCAGAACTTAAAGTAAC	386
Db	241	GACTGTGCTGCTTACTTTTAGTGTGATGATTCACAAAGCTTCCAGAACTTAAAGTAAC	300
Qy	387	TGCAAGAAAGAAATTCATATATTGACAGATGTGAAAGAGCCTGAGAGCTTCCCTTTGTG	446
Db	301	TGCAAGAAAGAAATTCATATATTGACAGATGTGAAAGAGCCTGAGAGCTTCCCTTTGTG	360
Qy	447	ATTCTGGTAAACAAGATTGACATAAGCGAACCGCAGGTGTCTACAGAAGAGCCCAAGCT	506
Db	361	ATTCTGGTAAACAAGATTGACATAAGCGAACCGCAGGTGTCTACAGAAGAGCCCAAGCT	420
Qy	507	TGCTGCAGGACCAACCGGACTATCTTTATTTTGAACAGTGCACAAAGATGCCAAGAT	566
Db	421	TGCTGCAGGACCAACCGGACTATCTTTATTTTGAACAGTGCACAAAGATGCCAAGAT	480
Qy	567	GTGCGACGACCTTTGAGGAAGCGGTTTCGAAGAGTTCTTGCTACCGAGGATAGTGCAGAT	626
Db	481	GTGCGACGACCTTTGAGGAAGCGGTTTCGAAGAGTTCTTGCTACCGAGGATAGTGCAGAT	540
Qy	627	CATTTCATTTCAGACACACAGTCAATCTTCCGAAAGCCCAAGCCTAGCTCATCTTGC	686
Db	541	CATTTCATTTCAGACACACAGTCAATCTTCCGAAAGCCCAAGCCTAGCTCATCTTGC	600
Qy	687	TGTTGA	692
Db	601	TGTTGA	606

RESULT 7
US-09-759-143-206/c
; Sequence 206, Application US/09759143
; Patent No. US2002022248A1

```

; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 206
; LENGTH: 487
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(487)
; OTHER INFORMATION: n = A,T,C or G
US-09-759-143-206

Query Match 37.4%; Score 439.6; DB 9; Length 487;
Best Local Similarity 96.0%; Pred. No. 1.7e-88;
Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

Qy 536 TTTTGAACAAGTGCACCAAGATGCCCAATGCGCAGACGACGCTTTGAGGAAGCGGTTCG 595
Db 487 TTTTGAACAAGTGCACCAAGATGCCCAATGCGCAGACGACGCTTTGAGGAAGCGGTTCG 428

Qy 596 AAGAGTTCCTTGCTACCGAGGATAGGTGAGATCATCTTTGATTTCAGACAGACACAGTCAATCT 655
Db 427 AAGAGTTCCTTGCTACCGAGGATAGGTGAGATCATCTTTGATTTCAGACAGACACAGTCAATCT 368

Qy 656 TCACCGAAGCCCAAGCCTAGC-TCATCTGCTGTTGATTGTTAGATTGTTGATGATTC 714
Db 367 TCACCGAAGCCCAAGCCTAGCTCATCTGCTGTTGATTGTTAGATTGTTGATGATTC 308

Qy 715 TAACCAACTCACAACATATACAAAAATCAACATGCGGATGGAAGAGAATTAGCGTTTG 774
Db 307 TNAACCAACTCACAACATATACAAATATACATGCGGATGGAAGAGAATTAGCGTTTG 248

Qy 775 CAGCAGTGATCATCTACTAATAAATAAATAAATGTTGCTGCTTCATTAGTTGGTGGG 834
Db 247 CAGCAGTGATCATCTACTAATAAATAAATAAATGTTGCTGCTTCATTAGTTGGTGGG 188

Qy 835 AGAAGGGACACATCCACTCTTCGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 894
Db 187 AGAAGGGACACATCCACTCTTCGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 128

Qy 895 TAAATTGTAACTGTGTCTTAATAACGTTTCTTTTAAATTTAAATATGTAAGTTGCAGAGCTA 954
Db 127 TAAATTGTAACTGTGTCTTAATAACGTTTCTTTTAAATTTAAATATGTAAGTTGCAGAGCTA 68

Qy 955 ATAATGAATGACCCAGACTTTAATATATAAATAAATAAAGAACTTGACTA 1006
Db 67 ATAATGAATGACCCAGACTTTAATATATAAATAAATAAAGAACTTGACTA 16

RESULT 8

```

US-09-780-669-206/c
 ; Sequence 206, Application US/09780669
 ; Patent No. US2002005197A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Mitcham, Jennifer L.
 ; APPLICANT: Harlocker, Susan L.
 ; APPLICANT: Jiang, Yuqi
 ; APPLICANT: Henderson, Robert A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fenger, Gary R.
 ; APPLICANT: Retter, Marc W.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Day, Craig H.
 ; APPLICANT: Vedvick, Thomas S.
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Li, Samuel
 ; APPLICANT: Wang, Aijun
 ; APPLICANT: Skelky, Yasir A.W.
 ; APPLICANT: Hepler, William
 ; APPLICANT: Hural, John
 ; APPLICANT: McNeill, Patricia D.
 ; APPLICANT: Houghton, Raymond L.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 ; FILE REFERENCES: 210121.427C24
 ; CURRENT APPLICATION NUMBER: US/09/780,669
 ; CURRENT FILING DATE: 2001-02-09
 ; NUMBER OF SEQ ID NOS: 943
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 206
 ; LENGTH: 487
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc.feature
 ; LOCATION: (1)...(487)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-780-669-206

Query Match 37.4%; Score 439.6; DB 9; Length 487;
 Best Local Similarity 96.0%; Pred. No. 1.7e-88;
 Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
 QY 536 TTTTGAACAAGTGCAAAAGATGCCCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTTTCG 595
 Db 487 TTTTGAACAAGTGCAAAAGATGCCCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTTTCG 428
 QY 596 AAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCT 655
 Db 427 AAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCT 368
 QY 656 TCACCGAAAGCCCAAGCCCTAGC-TCATCTTGTCTGTGATTGTGATTGTGATGCAATTC 714
 Db 367 TCACCGAAAGCCCAAGCCCTAGC-TCATCTTGTCTGTGATTGTGATTGTGATGCAATTC 308
 QY 715 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAAATAGCGTTTG 774
 Db 307 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAAATAGCGTTTG 248
 QY 775 CAGCAGTGTATCATCTACTTAATAAATTAACCTAATGTCTGCTTCAATTAGTTGGTGG 834
 Db 247 CAGCAGTGTATCATCTACTTAATAAATTAACCTAATGTCTGCTTCAATTAGTTGGTGG 188
 QY 835 AGAAGGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 894
 Db 187 AGAAGGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 128
 QY 895 TAAATTTGAACAGTTGTCTTAATAACGTTTCTTTAATTTAAATPATGTAAGTTGCAGAGCTA 954
 Db 127 TAAATTTGAACAGTTGTCTTAATAACGTTTCTTTAATTTAAATPATGTAAGTTGCAGAGCTA 68

QY 955 ATAAATGAATGACCAAGACTTTTAAATTAATAAATAAAGAACTTGACTA 1006
 Db 67 ATAAATGAATGACCAAGACTTTTAAATTAATAAATAAAGAACTTGACTA 16
 RESULT 9
 US-09-030-606-206/c
 ; Sequence 206, Application US/09030606
 ; Patent No. US20020081580A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jiangchun
 ; APPLICANT: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNODIAGNOSIS OF PROSTATE CANCER AND METHODS
 ; NUMBER OF SEQUENCES: 224
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/030,606
 ; FILING DATE: 25-FEB-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.428C3
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (206) 622-4900
 ; TELEFAX: (206) 682-6031
 ; INFORMATION FOR SEQ ID NO: 206:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 487 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 US-09-030-606-206

Query Match 37.4%; Score 439.6; DB 9; Length 487;
 Best Local Similarity 96.0%; Pred. No. 1.7e-88;
 Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;
 QY 536 TTTTGAACAAGTGCAAAAGATGCCCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTTTCG 595
 Db 487 TTTTGAACAAGTGCAAAAGATGCCCAAAATGTGGCAGAGCGCTTTGAGGAAGCGTTTCG 428
 QY 596 AAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCT 655
 Db 427 AAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGATTTCAGACAGACACAGTCAATCT 368
 QY 656 TCACCGAAAGCCCAAGCCCTAGC-TCATCTTGTCTGTGATTGTGATTGTGATGCAATTC 714
 Db 367 TCACCGAAAGCCCAAGCCCTAGC-TCATCTTGTCTGTGATTGTGATTGTGATGCAATTC 308
 QY 715 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAAATAGCGTTTG 774
 Db 307 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAAATAGCGTTTG 248
 QY 775 CAGCAGTGTATCATCTACTTAATAAATTAACCTAATGTCTGCTTCAATTAGTTGGTGG 834
 Db 247 CAGCAGTGTATCATCTACTTAATAAATTAACCTAATGTCTGCTTCAATTAGTTGGTGG 188
 QY 835 AGAAGGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 894
 Db 187 AGAAGGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGCACCTTACATTTA 128

APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Hepler, William T.
APPLICANT: Henderson, Robert A.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals de Bassols, Carlota
APPLICANT: Foy, Teresa
APPLICANT: Fanger, Gary R.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C26
CURRENT APPLICATION NUMBER: US/09/895,814
CURRENT FILING DATE: 2001-06-29
NUMBER OF SEQ ID NOS: 990
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 206
LENGTH: 487
TYPE: DNA
ORGANISM: Homo sapien
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)-(487)
OTHER INFORMATION: n = A,T,C or G
US-09-895-814-206

Query Match 37.4%; Score 439.6; DB 10; Length 487;
Best Local Similarity 96.0%; Pred. No. 1.7e-88;
Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 536 TTTTGAACAAGTGCAAAAGATGCCAAATGTGGCAGAGCGCTTTGAGGAAGCGGTTCG 595
DB 487 TTTTGAACAAGTGCAAAAGATGCCAAATGTGGCAGAGCGCTTTGAGGAAGCGGTTCG 428
QY 596 AAGAGTTCTTGTACCGAGGATAGGTGAGTTCAGATCATTTGATTCAGACAGACAGTCAATCT 655
DB 427 AAGAGTTCTTGTACCGAGGATAGGTGAGTTCAGATCATTTGATTCAGACAGACAGTCTATTT 368
QY 715 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAGAAATTTAGCGTTTG 774
DB 307 TNACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAGAAATTTAGCGTTTG 248
QY 775 CAGCAGTGTATCATCTACTAATAAATTAATACTAATGTTGCTGCTTCATTAGTTGGTGG 834
DB 247 CAGCAGTGTATCATCTACTAATAAATTAATACTAATGTTGCTGCTTCATTAGTTGGTGG 188
QY 835 AGAAGGGACACATCCACTCTTGGAGAAATATATTTACTCAATAATGGCACCTTACATTTA 894
DB 187 AGAAGGGACACATCCACTCTTGGAGAAATATATTTACTCAATAATGGCACCTTACATTTA 128
QY 895 TAAATTTGAACAGTTGCTTAATAACCTTTCTTTTAAATTAATTAATGTTGAGAGCTA 954
DB 127 TAAATTTGAACAGTTGCTTAATAACCTTTCTTTTAAATTAATTAATGTTGAGAGCTA 68
QY 955 ATAAATGAATGACCAAGCTTTAATTAATAAATAAATAAAGAACTTGACTA 1006
DB 67 ATAAATGAATGACCAAGCTTTAATTAATAAATAAATAAAGAACTTGACTA 16

RESULT 15

US-10-144-678A-206/C
Sequence 206, Application US/10144678A
Publication No. US20030157089A1
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan L.
APPLICANT: Jiang, Yugu

APPLICANT: Henderson, Robert A.
APPLICANT: Kalos, Michael D.
APPLICANT: Fanger, Gary R.
APPLICANT: Retter, Marc W.
APPLICANT: Stolk, John A.
APPLICANT: Day, Craig H.
APPLICANT: Vedick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Li, Samuel X.
APPLICANT: Wang, Aljun
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Hepler, William T.
APPLICANT: Hural, John
APPLICANT: McNeill, Patricia D.
APPLICANT: Houghton, Raymond L.
APPLICANT: Vinals y de Bassols, Carlota
APPLICANT: Foy, Teresa M.
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Deng, Ta
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C28
CURRENT APPLICATION NUMBER: US/10/144,678A
CURRENT FILING DATE: 2002-08-12
NUMBER OF SEQ ID NOS: 1033
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 206
LENGTH: 487
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 29, 106, 119, 154, 204, 236, 279, 306, 318, 325, 353, 448,
LOCATION: 466, 470
OTHER INFORMATION: n = A,T,C or G
US-10-144-678A-206

Query Match 37.4%; Score 439.6; DB 13; Length 487;
Best Local Similarity 96.0%; Pred. No. 1.7e-88;
Matches 453; Conservative 0; Mismatches 18; Indels 1; Gaps 1;

QY 536 TTTTGAACAAGTGCAAAAGATGCCAAATGTGGCAGAGCGCTTTGAGGAAGCGGTTCG 595
DB 487 TTTTGAACAAGTGCAAAAGATGCCAAATGTGGCAGAGCGCTTTGAGGAAGCGGTTCG 428
QY 596 AAGAGTTCTTGTACCGAGGATAGGTGAGTTCAGATCATTTGATTCAGACAGACAGTCAATCT 655
DB 427 AAGAGTTCTTGTACCGAGGATAGGTGAGTTCAGATCATTTGATTCAGACAGACAGTCTATTT 368
QY 656 TCACCGAAGCCCAAGCGCTAGC-TCATCTTGTCTGTTGATTGTAGATTGTTGATGCTTTC 714
DB 367 TCACCGAAGCCCAAGCGCTAGC-TCATCTTGTCTGTTGATTGTAGATTGTTGATGCTTTC 308
QY 715 TAACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAGAAATTTAGCGTTTG 774
DB 307 TNACCAACTCACACATATACAAAATCAACATGGGATGGAGAGAGAAATTTAGCGTTTG 248
QY 775 CAGCAGTGTATCATCTACTAATAAATTAATACTAATGTTGCTGCTTCATTAGTTGGTGG 834
DB 247 CAGCAGTGTATCATCTACTAATAAATTAATACTAATGTTGCTGCTTCATTAGTTGGTGG 188
QY 835 AGAAGGGACACATCCACTCTTGGAGAAATATATTTACTCAATAATGGCACCTTACATTTA 894
DB 187 AGAAGGGACACATCCACTCTTGGAGAAATATATTTACTCAATAATGGCACCTTACATTTA 128
QY 895 TAAATTTGAACAGTTGCTTAATAACCTTTCTTTTAAATTAATTAATGTTGAGAGCTA 954
DB 127 TAAATTTGAACAGTTGCTTAATAACCTTTCTTTTAAATTAATTAATGTTGAGAGCTA 68
QY 955 ATAAATGAATGACCAAGCTTTAATTAATAAATAAATAAAGAACTTGACTA 1006
DB 67 ATAAATGAATGACCAAGCTTTAATTAATAAATAAATAAAGAACTTGACTA 16

Thu Feb 19 09:29:46 2004

us-09-988-974-6.rnpb

Page 11

Search completed: February 10, 2004, 04:54:25
Job time : 494 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2004, 00:45:41 ; Search time 2885 Seconds

(without alignments)
9898.703 Million cell updates/sec

Title: US-09-988-974-6

Perfect score: 1175

Sequence: 1 CACTGTGATGAAACACTTTT.....GTAAAAA.....1175

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estm:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pin:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	888	75.6	1201	13	BX422913
2	832.4	70.8	1201	13	BX422914
3	808.2	68.8	1201	13	BX361467
4	802.2	68.3	837	13	BQ221697

5	790.2	67.3	835	12	BI761520
6	720.4	61.3	791	9	AI671547
7	713.2	60.7	781	9	AI934576
8	712	60.6	734	9	AI833027
9	707.6	60.2	728	12	BM981745
10	706.6	60.1	1027	13	BM981745
11	704.4	59.9	720	13	BQ772388
12	704	59.9	881	14	CD515129
13	692.2	58.9	767	9	AI479232
14	688.6	58.6	778	9	AI758228
15	686	58.4	708	9	AI758228
16	685.6	58.3	714	14	CA308701
17	679.8	57.9	696	9	AI659875
18	675	57.4	715	9	AI763029
19	667.4	56.8	705	10	EG054716
20	664.8	56.6	734	9	AI765395
21	659.4	56.1	678	13	BU682548
22	659.2	56.1	1205	11	AK032133
23	659.2	56.1	1242	11	AK017301
24	659	56.1	659	14	CA441619
25	653.8	55.6	1072	12	BM476930
26	635.6	54.1	682	10	BE326331
27	617.6	52.6	823	14	CA488998
28	617	52.5	958	10	BE858526
29	616.4	52.5	910	12	BI825651
30	609	51.8	609	12	BM353471
31	605.6	51.5	613	9	AI922378
32	604.4	51.4	656	9	AI423136
33	593	50.5	887	13	EX352178
34	592	50.4	720	12	BI818653
35	589.4	50.2	614	12	BM674469
36	588	50.0	609	12	BM717088
37	585	49.8	615	13	BU616781
38	583.4	49.7	616	12	BG874452
39	582.8	49.6	1179	12	BM926730
40	575.4	49.0	588	9	AW614366
41	566.8	48.2	576	12	BM126004
42	566.6	48.2	761	12	BI766528
43	565.2	48.1	571	9	AI281626
44	561	47.7	561	9	AA563671
45	560.8	47.7	573	10	BE464245

ALIGNMENTS

RESULT 1
BX422913/C
LOCUS BX422913 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CSODN002YB16 3-PRIME, mRNA sequence.
ACCESSION BX422913
VERSION BX422913.1 GI:30655413
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 11038.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODN002YB16&cluster=11038.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DN002DA08NP1.

FEATURES
source
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN002YB16"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT 323 a 243 c 226 g 320 t 89 others
ORIGIN
Query Match 75.6%; Score 888; DB 13; Length 1201;
Best Local Similarity 97.3%; Pred. No. 2.6e-111;
Matches 891; Conservative 8; Mismatches 17; Indels 0; Gaps 0;
QY 58 GAGGGTCTTGAAGCTTTGAGATTAAACAATGCGAGGAAATCATCACTTTTAAAGTAA 117
DB 916 GAGGGTCTTGAAGCTTTGAGATTAAACAATGCGAGGAAATCATCACTTTTAAAGTAA 857
QY 118 TTCTCTCTGGAGTGTGGAGTTGGAAGATTCATTATGAACAGATATGTAACATAA 177
DB 856 TTCTCTCTGGAGTGTGGAGTTGGAAGATTCATTATGAACAGATATGTAACATAA 797
QY 178 AGTTTCATACCCAGCTCTTCCATACATAGGTGTGGAATTTTAAATAAGATTGGAG 237
DB 796 AGTTTCATACCCAGCTCTTCCATACATAGGTGTGGAATTTTAAATAAGATTGGAG 737
QY 238 TGATGGAATTTGTTACATGCAATTTGGACACCGCAGGTGAGAGGATTCGAA 297
DB 736 TGGATGGACATTTTGTACATGCAATTTGGACACCGCAGGTGAGAGGATTCGAA 677
QY 298 GCTGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTAGTGTGCGATG 357
DB 676 GCTGAGGACACATTTTACAGAGTTCTGACTGCTGCTGCTTACTTTAGTGTGCGATG 617
QY 358 ATTCAAAAGCTTCCAGAACTTAAAGTAATCTGGAAGAAAGATTCATATATTATGCGATG 417
DB 616 ATTCAAAAGCTTCCAGAACTTAAAGTAATCTGGAAGAAAGATTCATATATTATGCGATG 557
QY 418 TGAAGAGCTGAGAGCTTTCTTTTCTGATTTCTGGTAAACAAGATTGACATGCGAAC 477
DB 556 TGAAGAGCTGAGAGCTTTCTTTTCTGATTTCTGGTAAACAAGATTTACATGCGAAC 497
QY 478 GGCAGGTGCTACAGAAAGCCCAAGCTTGGTGAGGACAAACGCGCACTATCCTTATT 537
DB 496 GGCAGGTGCTACAGAAAGCCCAAGCTTGGTGAGGACAAACGCGCACTATCCTTATT 437
QY 538 TTGAACAAGTGCACAAAGATGCAAAATGTGGCAGCAGCCCTTTGAGGAGCGGTTCCAA 597
DB 436 TTGAACAAGTGCACAAAGATGCAAAATGTGGCAGCAGCCCTTTGAGGAGCGGTTCCAA 377
QY 598 GAGTTCTTCTACGAGGATAGTCAATCTGATTCATTCAGACACACAGTCATCTTCT 657
DB 376 GAGTTCTTCTACGAGGATAGTCAATCTGATTCATTCAGACACACAGTCATCTTCT 317
QY 658 ACCGAAGCCCAAGCTTAGCTCATCTTGTGTTGATTTGATTTGATTTGATTTGATTT 717
DB 316 ACCGAAGCCCAAGCTTAGCTCATCTTGTGTTGATTTGATTTGATTTGATTTGATTT 257
QY 718 CCACTCACAATATACAAATACACATGGGATGGAGAGAAATAGCGTTGCGAG 777
DB 256 CCAATTAACAATATACAAATACACATGGGATGGAGAGAAATAGCGTTGCGAG 197
QY 778 CAGTGTATCATCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 837
DB 196 CAGTGTATCATCTACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 137

QY 838 AGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGCGACCTTACATTATAA 897
DB 136 AGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGCGACCTTACATTATAA 77
QY 898 ATTGTAACAGTGTCTTAATAACCTTTCTTTAAATTAATGTAAGTTGAGAGCTAATA 957
DB 76 ATTGTAACAGTGTCTTAATAACCTTTCTTTAAATTAATGTAAGTTGAGAGCTAATA 17
QY 958 AATGAATGACCAAGA 973
DB 16 AATGAATGACCAAGA 1
RESULT 2
BX422914 1201 bp mRNA linear EST 13-MAY-2003
LOCUS BX422914 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS0DN002YB16 5-PRIME, mRNA sequence.
ACCESSION BX422914
VERSION BX422914.1 GI:30659414
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 1201)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 131 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 11038.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DN002DA08NP1&cluster=11038.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DN002DA08NP1.
Location/Qualifiers
1. .1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN002YB16"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."
BASE COUNT 357 a 214 c 253 g 308 t 59 others
ORIGIN
Query Match 70.8%; Score 832.4; DB 13; Length 1201;
Best Local Similarity 93.8%; Pred. No. 8.8e-104;
Matches 884; Conservative 12; Mismatches 43; Indels 3; Gaps 3;
QY 58 GAGGGTCTTGAAGCTTTGAGATTAAACAATGCGAGGAAATCATCACTTTTAAAGTAA 117
DB 159 GAGGGTCTTGAAGCTTTGAGATTAAACAATGCGAGGAAATCATCACTTTTAAAGTAA 218
QY 118 TTCTCTCTGGAGTGTGGAGTTGGGAGAGTTCATTATGAACAGATATGTAACATAA 177
DB 219 TTCTCTCTGGAGTGTGGAGTTGGGAGAGTTCATTATGAACAGATATGTAACATAA 278
QY 178 AGTTGATACCCAGCTCTTCCATACATAGGTGTGGAATTTTAAATAAGATTTCGAAG 237

DNA Sequencing by: Inocyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LRAM1466 row: n column: 15
 High quality sequence stop: 835.

FEATURES
 Location/Qualifiers
 1..835

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:5186894"
 /lab_host="DH10B"
 /clone_lib="NIH MGC 116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 PCMV-SPORE6; Site 1: Not I; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH MGC Library."
 BASE COUNT 269 a 145 c 186 g 235 t
 ORIGIN

Query Match 67.3%; Score 790.2; DB 12; Length 835;
 Best Local Similarity 98.8%; Pred. No. 5.5e-98;
 Matches 81; Conservative 0; Mismatches 8; Indels 2; Gaps 2;

QY 231 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 290
 DB 2 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 60
 QY 291 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 350
 DB 61 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 120
 QY 351 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 410
 DB 121 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 180
 QY 411 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 470
 DB 181 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 240
 QY 471 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 530
 DB 241 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 300
 QY 531 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 590
 DB 301 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 360
 QY 591 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 650
 DB 361 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 420
 QY 651 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 710
 DB 421 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 480
 QY 711 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 770
 DB 481 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 540
 QY 771 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 830
 DB 541 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 600
 QY 831 TTGGAAGTGGATGACATTTTGTACCATCGATTGGGACACGGCAGGTGAGGACGA 890

Db 601 TGGGAGAGGGACACATCCACTCTTGGAGGAATATATTTACTCAATAATGGCACCTTACA 660
 QY 891 TTTATAAATTTGACAGTTGCTATAAAGCTTTCTTTAATTTAAATATGTAAGTTGCAGA 950
 Db 661 TTTATAAATTTGACAGTTGCTATAAAGCTTTCTTTAATTTAAATATGTAAGTTGCAGA 720
 QY 951 GCTAATAAATGAATGACCAAGACTTTAA-TTATAATAAATAAGAAACTTGACTATTC 1009
 Db 721 GCTAATAAATGAATGACCAAGACTTTAAAGTCATAATAAATAAGAAACTTGACTATTC 780
 QY 1010 TAGAAGTTATATCTGCTGATTTTCTCTGGGAAAATGGGAACTACTCTTT 1056
 Db 781 TAGAAGTTATATCTGCTGATTTTCTCTGGGAAAATGGGAACTACTCTTT 827

RESULT 6

AI671547/c

LOCUS

DEFINITION

AI671547

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AI671547 791 bp mRNA linear EST 17-DEC-1999
 wb33e07.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2307492 3'
 similar to SW:RAB9_HUMAN P51151 RAS-RELATED PROTEIN RAB-9. ; mRNA
 sequence.
 AI671547
 EST
 AI671547.1 GI:4851278
 Homo sapiens (human)
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 791)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Insert Length: 536 Std Error: 0.00
 Seq primer: -40UP from Gibco
 High quality sequence stop: 439.

Location/Qualifiers
 1..791
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2307492"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP GC6"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
 from the normalized library NCI-CGAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification, and
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clonesDs
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 214 a 170 c 139 g 267 t 1 others

FEATURES

source

1..791

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2307492"

/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"

/clone_lib="NCI CGAP GC6"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA

from the normalized library NCI-CGAP GC4 was prepared, and

ss circles were made in vitro. Following HAP purification, and

this DNA was used as tracer in a subtractive hybridization

reaction. The driver was PCR-amplified cDNAs from a pool

of 5,000 clones made from the same library (clonesDs

1257096-1258631, 1469064-1470983, and 1475592-1476743).

Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 214 a 170 c 139 g 267 t 1 others

ORIGIN

Query Match 61.3%; Score 720.4; DB 9; Length 791;

Best Local Similarity 98.1%; Pred. No. 1.6e-88;

Matches 760; Conservative 0; Mismatches 12; Indels 3; Gaps 3;

QY 229 ATTGGAGTGGATGGACATTTGTTTACCATGACGATTTGGGACACGGCAGGTCCAGAGC 288
 Db |||||
 QY 790 ATTGGAGTGGATGGAGCA-TTTGTTACCGAGGATTT-GGACACGGCAGGTCCAGAGC 733
 Db |||||
 QY 289 GATTCGAGGCTGAGGACACCAATTTTACAGAGGTTCTGACTCTGCTGCTTACTTTTA 348
 Db |||||
 QY 732 GATCCGAGGCTGAGGACA-CATTTTACAGAGTTTCTGACTCTGCTGCTTACTTTTA 674
 Db |||||
 QY 349 GTGTGATGATTCACAAAGCTTCCAGACTTAAGTAACTGGAAGAAGAAATTCATATTT 408
 Db |||||
 QY 673 GTGTGATGATTCACAAAGCTTCCAGACTTAAGTAACTGGAAGAAGAAATTCATATTT 614
 Db |||||
 QY 409 ATGCAGATGTCAAGAGCCTGAGAGCTTCTCTTTGTGATTCCTGGGTAAACAAGATTGACA 468
 Db |||||
 QY 613 ATGCAGATGTCAAGAGCCTGAGAGCTTCTCTTTGTGATTCCTGGGTAAACAAGATTGACA 554
 Db |||||
 QY 469 TAAGCGAAGCGAGGTGTCTACAGAGAAGCCCAAGCTTGGTGCAGGGCAACGGCGACT 528
 Db |||||
 QY 553 TAAGCGAAGCGAGGTGTCTACAGAGAAGCCCAAGCTTGGTGCAGGGCAACGGCGACT 494
 Db |||||
 QY 529 ATCTCTATTTTGAACAAGTGCACAAAGATGCCCAATGTGGCAGCAGCTTTGGGAAG 588
 Db |||||
 QY 493 ATCTCTATTTTGAACAAGTGCACAAAGATGCCCAATGTGGCAGCAGCTTTGGGAAG 434
 Db |||||
 QY 589 CGGTTCGAGAGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACACACAG 648
 Db |||||
 QY 433 CGGTTCGAGAGTTCTTGTCTACGAGGATAGGTGAGATCATTTGATTCAGACACACAG 374
 Db |||||
 QY 649 TCATCTTCCAGAAAGCCCAAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
 Db |||||
 QY 373 TCATCTTCCAGAAAGCCCAAGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 314
 Db |||||
 QY 709 GCATTTCAACCACTCACACATATACACAAATCAACATGGGATGGAGAGAGATTAAG 768
 Db |||||
 QY 313 GCATTTCAACCACTCACACATATACACAAATCAACATGGGATGGAGAGAGATTAAG 254
 Db |||||
 QY 769 CGTTTCGAGAGTGTATCTACTATAATAATAATAATAATAATAATAATAATAATAATA 828
 Db |||||
 QY 253 CGTTTCGAGAGTGTATCTACTATAATAATAATAATAATAATAATAATAATAATAATA 194
 Db |||||
 QY 829 GGTGGAGAGGGACACATCCATCTCTGGAGGATATATTTACTCAATATGACACCTTA 888
 Db |||||
 QY 193 GGTGGAGAGGGACACATCCATCTCTGGAGGATATATTTACTCAATATGACACCTTA 134
 Db |||||
 QY 889 CATTTATAATGTAACAGTTGTCTTAATACGTTTCTTTAAATTAATTAATTAATTAAT 948
 Db |||||
 QY 133 CATTTATAATGTAACAGTTGTCTTAATACGTTTCTTTAAATTAATTAATTAATTAAT 74
 Db |||||
 QY 949 GAGCTAATAATGAATGACCAAGACTTTTAATTAATAATAATAATAATAATAATAATA 1003
 Db |||||
 QY 73 GAGCTAATAATGAATGACCAAGACTTTTAATTAATAATAATAATAATAATAATAATA 19

RESULT 7
 AI934576/c
 LOCUS 781 bp mRNA linear EST 02-SEP-1999
 DEFINITION wp09f06.x1 NCI CGAP kid12 Homo sapiens cDNA clone IMAGE:2464355 3', similar to SW:RAB9_HUMAN P5151 RAS-RELATED PROTEIN RAB-9.; mRNA sequence.
 ACCESSION AI934576
 VERSION AI934576.1 GI:5673373
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 781)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-1@mail.nih.gov

Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 www-bio.lnlni.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 469.

FEATURES

source
 1..781
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2464355"
 /tissue_type="2 pooled tumors (clear cell type)"
 /lab_host="DH10B"
 /clone_lib="NCI CGAP Kid12"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP_Kids was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1323912-1325831, 1471368-1472903 and 1492104-1493255). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 235 a 155 c 122 g 263 t 6 others
 ORIGIN
 Query Match 60.7%; Score 713.2; DB 9; Length 781;
 Best Local Similarity 96.7%; Pred. No. 1.5e-87;
 Matches 757; Conservative 0; Mismatches 23; Indels 3; Gaps 3;
 QY 390 AAGAAAGATTTCATATATTATGACAGATGTGAAAGAGCCTGAGAGCTTTCCTTTGTGATT 449
 Db |||||
 QY 781 AAGAAAGATTTCATATATTATGACAGATGTGAAAGAGCCTGAGAGCTTTCCTTTGTGATT 722
 Db |||||
 QY 450 CTGGTTCACAGATTGACATAAGCGACGGCAGGTGTACAGAG-AAGCCCAAGCTTG 508
 Db |||||
 QY 721 CTGGGT-ACAGATTGACATAAGCGACGGCAGGTGTACAGAGNAAGCCCAAGCTTG 663
 Db |||||
 QY 509 GTGCAGGACACCGCGACTATCCTTATTTTGAACAAAGTGCAAAAGATGCCCAAAATGT 568
 Db |||||
 QY 662 GTGCAGGACACACCGCGACTATCCTTATTTG-ACCAAGTGCAAAAGATGCCCAAAATGT 604
 Db |||||
 QY 569 GGCAGCAGCCTTTGAGGAAGGGTTCGAGAGTTCCTGCTACCGAGGATAGGTCAATCA 628
 Db |||||
 QY 603 GGCAGCAGCCTTTGAGGAAGGGTTCGAGAGTTCCTGCTACCGAGGATAGGTCAATCA 544
 Db |||||
 QY 629 TTTGATTTCAGACACACAGTCAATCTTCACGAAAGCCCAAGCCTAGCTCATCTTCTGCTG 688
 Db |||||
 QY 543 TTTGATTTCAGACACACAGTCAATCTTCACGAAAGCCCAAGCCTAGCTCATCTTCTGCTG 484
 Db |||||
 QY 589 TTGATTGTTAGATTGTTGATGCAATTTTACCAACTCACATATACACAAAATCAACATG 748
 Db |||||
 QY 483 TTGATTGTTAGATTGTTGATGCAATTTTACCAACTCACATATACACAAAATCAACATG 424
 Db |||||
 QY 749 GGGATGGAAGAAGAAATTAGCGTTTGCGAGCAGTGTATCATCTACTAATAAATAAATAA 808
 Db |||||
 QY 423 GGGATGGAAGAAGAAATTAGCGTTTGCGAGCAGTGTATCATCTACTAATAAATAAATAA 364
 Db |||||
 QY 809 ATGTTGCTGCTTCAATTAGTTGGTGGGAAGAGGACACATCCACTCTTTGGAGGAATATTT 868
 Db |||||
 QY 363 ATGTTGCTGCTTCAATTAGTTGGTGGGAAGAGACACATCCACTCTTTGGAGGAATATTT 304
 Db |||||
 QY 869 TACTCAATATGCAACCTTACATTTAATTTGTAACAGTTGCTAATAACGTTTCTTTTA 928
 Db |||||
 QY 303 TACTCAATATGCAACCTTACATTTAATTTGTAACAGTTGCTAATAACGTTTCTTTTA 244
 Db |||||
 QY 929 ATTTAAATATGTAAGTTGCAGAGCTAATAAATGAAATGACCAAGACTTTTAAATTAATA 988

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library prepared by: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 1-48, >AT rich#Low_complexity (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes

FEATURES

source
1. .728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-EN1-adi-p-23-0-UI"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial Cells"
/dev_stage="Adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-CF-EN1"
/note="Organ: Lung; Vector: pVT3-Pac (Pharmacia) with a modified Ogan linker; Site 1: EcoR I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldi, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)18 tail. The sequence tag for this library is CTGCTCAGGT.
TAG LIB=UI-CF-EN1
TAG_TISSUE=Human Lung Epithelial Cell Lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTGCTCAGGT"
BASE COUNT 220 a 140 c 111 g 257 t
ORIGIN

Query Match 60.2%; Score 707.6; DB 12; Length 728;
Best Local Similarity 98.8%; Pred. No. 9.1e-87;
Matches 713; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 454 GTAACAAGATTGACATACGAGCGAGGTCTCTACAGAGAGCCCAAGCTTGTCGA 513
DB 728 GTAACAAGATTGACATACGAGCGAGGTCTCTACAGAGAGCCCAAGCTTGTCGA 669
QY 514 GGGACACCGCGACTATCTTTTGAACAGTGCAGAGATGCCCAATGTGGCAG 573
DB 668 GGGACACCGCGACTATCTTTTGAACAGTGCAGAGATGCCCAATGTGGCAG 609
QY 574 CAGCCTTTGAGGAAGCGGTTCGAAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGA 633
DB 608 CAGCCTTTGAGGAAGCGGTTCGAAGAGTTCTTGTCTACCGAGGATAGGTTCAGATCATTTGA 549
QY 634 TTTCAGACACACAGTCAATCTTCCAGAACCCCAAGCTAGCTCATCTTCTGCTTGAT 693
DB 548 TTTCAGACACACAGTCAATCTTCCAGAACCCCAAGCTAGCTCATCTTCTGCTTGAT 489

QY 694 TGTTAGATTCTTGATGCATTTCTAACCAACTCACAATATACACATAATCAACATGGGGAT 753
DB 488 TGTTAGATTCTTGATGCATTTCTAACCAACTCACAATATACACATAATCAACATGGGGAT 429
QY 754 GGGAGAGAGAAATTAGCGTTTTCAGAGAGTGTATCTACTACTATAATAATTAACATAATGTT 813
DB 428 GGGAGAGAGAAATTAGCGTTTTCAGAGAGTGTATCTACTACTATAATAATTAACATAATGTT 369
QY 814 GCTGCTTCATTAGTTTGGTGGGAGAGGGACACATCCACTCTTTCGGAGGAATATATTTTACTC 873
DB 368 GCTGCTTCATTAGTTTGGTGGGAGAGGGACACATCCACTCTTTCGGAGGAATATATTTTACTC 309
QY 874 AATAATGGCACCTTACATTTTATAAATTTGAACAGTTGTCTAATAAGCTTCTTTAATTTA 933
DB 308 AATAATGGCACCTTACATTTTATAAATTTGAACAGTTGTCTAATAAGCTTCTTTAATTTA 249
QY 934 AATATGTAAGTTCAGAGCTTAATAAATGAAATCACCAGACCTTTTAATTAATAAAAAATA 993
DB 248 AATATGTAAGTTCAGAGCTTAATAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 189
QY 994 AGAACTTGACTATTTTAGAAGTTATACITGGATTTTTTCTGGGAAATGGAGAACTAC 1053
DB 188 AGAACTTGACTATTTTAGAAGTTATACITGGATTTTTTCTGGGAAATGGAGAACTAC 129
QY 1054 TTTTATATGTTGATGTTTTTATGCAATTAGCATTTGTTCTTGGTTCAGGGAATACCTT 1113
DB 128 TTTTATATGTTGATGTTTTTATGCAATTAGCATTTGTTCTTGGTTCAGGGAATACCTT 69
QY 1114 TCCTAAAGCAATTAATGTTAGATATTAAAGATTAAAAATCTAATGTAATAAAAAAATA 1173
DB 68 TCCTAAAGCAATTAATGTTAGATATTAAAGATTAAAAATCTAATGTAATAAAAAAATA 9
QY 1174 AA 1175
DB 8 AA 7
RESULT 10
BX344135
LOCUS
DEFINITION BX344135 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
Clone CS0D1060YA02 5-PRIME, mRNA sequence.
ACCESSION BX344135
VERSION BX344135.1 GI:30338244
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1027)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 11038.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1A1015ZG02QP1&cluster=11038.f. Contact :
Peng Liang Email : filiang@life.techn.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS1A1015ZG02QP1.
FEATURES
source
1. 1027
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1060YA02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

```

BASE COUNT      267 a   222 c   193 g   289 t   56 others
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

BASE COUNT	267 a	222 c	193 g	289 t	56 others
ORIGIN					
Query Match			60.1%	Score 706.6;	DB 13; Length 1027;
Best Local Similarity			89.1%	Pred. No. 1e-86;	
Matches 698; Conservative			50;	Mismatches 35;	Indels 0; Gaps 0

QY	1	CACTGTGATGAACAACCTTTTCCGGTGTGCTTTGAGTGCATCTTCTCACAACACCCCTTAGGAG	60
Db	245	CACTKTATTAACAACACTTTTCCGGTGTGCTTTGAKTTCACTTCTCAACAACCCCTAKGAK	304
QY	61	GGTCTTTGAAGCTTTTGAGATTAAACAATGGCAGGAAAATCATCACATTTTTTAAAGTAATTC	120
Db	305	GGTCTTTAAKCTTTTATATTAACAATKKCAKKAANAATCATCACATTTTTTAAATTAATTC	364
QY	121	TCCTTGGAGATGGTGGAGTTGGGAGAGTTCACTTATGAACAGATGTAACTTAATAAGT	180
Db	365	TCCTTKKATATKTKTKAKXTTTKGAAKATTTCACTTATKAACAATATKAACCTWAFAAT	424
QY	181	TTGATACCCAGCTCTTCCATACAATAGGTGTGSAATTTTTAAATAAAGATTGGGAATGG	240
Db	425	TTTATACCCAKCTCTTCCATACAATAGKTKTGTGAATTTTTAAATAAATATTTTKWATTTK	484
QY	241	ATGGCACATTTTCTAACCATGACAGATTTTGGGACACGGCAGCTCAGGAGCGAATCCGAAGCC	300
Db	485	WTTKACATTTTAAACCATKOWTATTTKKKACACKGCATKTCATKAGCTWTTTCGNAAGCC	544
QY	301	TGAGGACACCATTTTACAGAGTTCTGACTGTGCTCTGCTTACTTTTAGTGTGATGATT	360
Db	545	TGATKACACCATTTTACAGAGTTCGACTGTGCTCTGCTTACTTTTAGTGTGATGATT	604
QY	361	CACAAGCTTCCAGAACCTTAAGTAACCTGGAGAAAGAAATTCATATATTTATGCGAGATGGA	420
Db	605	CACAAGCTTCCAGAACCTTAAGTAACCTGGAGAAAGAAATTCATATATTTATGCGAGATGGA	664
QY	421	AAGAGCCTGAGAGCTTTCCTTTTGTGATTTCTGGGTAAACAAGATTGACATAAGCGAACGCG	480
Db	665	AAGAGCCTGARAGCTTTCCTTTTGTGATTTCTGGGTAAACAAGATTGACATAAGCGAACGCG	724
QY	481	AGGTGCTTACAGAAGAAGCCCAAGCTTGTGTGACGGGACAAAGCGGACTATTCCTTATTTTG	540
Db	725	AGGTGCTTACAGAAGAAGCCCAAGCTTGTGTGACGGGACAAAGCGGACTATTCCTTATTTTG	784
QY	541	AAAACAAGTGCAAAAGATGCCAATAATGTGGCAGCAGCGCTTTGAGGAAGCGGTTCCGAAG	600
Db	785	AAAACAAGTTCAAAAGATGCCAATAATGTGGCAGCAGCGCTTTGAGGAAGCGGTTCCGAAG	844
QY	601	TTCTTGTACCGAGGATAGGTCAGATCATTTTGATTCAGACAGACACAGTCAATCTTCACC	660
Db	845	TTCTTGTACCGAGGATAGGTCAGATCATTTTGATTCAGACAGACACAGTCAATCTTCACC	904
QY	661	GAAGGCCCAAGCCTAGCTCATCTTGTGTTGATTGTTAGATTGTTGATGCTTCAACCA	720
Db	905	GAAGGCCCAAGCCTAGCTCATCTTGTGTTGATTGTTAGATTGTTGATGCTTCAACCA	964
QY	721	ACTCACACATATACAAAAATCAACATGGGGATGGGAAGAGAAATTTAGCGTTTGAGGAG	780
Db	965	ACTCACACATATACACCAATCAACATGGGGATGGGAAGAGATTTAGCGTTTGAGGAG	1024
QY	781	TGT 783	
Db	1025	TVT 1027	

RESULT 11	BQ772388/c	LOCUS	BQ772388	720 bp	mRNA	linear	EST 26-JUL-2003
		DEFINITION	UI-H-FEO-bbq-e-19-0-UI.s1	NCI	CGAP_FEO	Homo sapiens	cDNA clone
			UI-H-FEO-bbq-e-19-0-UI 3'				mRNA sequence.

BQ772388
 BQ772388.1 GI:21980864
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 (1 bases 1 to 720)
 NC1-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapsb-remail@nih.gov
 Tissue Procurement: James Martin
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 cDNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be obtained
 from Dr. M. Bento Soares, bento-soares@uiowa.edu
 Seq primer: M13 FORWARD
 POLYA=yes.

```

FEATURES             Location/Qualifiers
     1..720
     /organism="Homo sapiens"
     /mol_type="mRNA"
     /db_xref="taxon:9606"
     /clone="UI-H-FE0-bbq-e-19-0-UI"
     /issue_type="Chondrosarcoma Cell line"
     /dev_stage="Adult"
     /lab_host="DH10B (Life Technologies)"
     /clone_lib="NCI CGAP FE0"
     /note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site_1: EcoR I; Site 2: Not I; NCI CGAP FE0 is
a cDNA library containing the following tissue(s): a pool
of 3 chondrosarcoma cell lines ( grade 2) The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA,
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCTACGGAC. The cell lines was provided by Dr James Martin
of University of Iowa.

```

	University of Iowa.
TAG_LIB=UI-H-FEO	
TAG_TISSUE=Human grade 2 chondrosarcoma cell line pool	
TAG_SEQ=CGCTACGCAC"	
216 a 138 c 112 g 254 t	
BASE COUNT	

Query Match	59.9%	Score 704.4	DB 13	Length 720
Best Local Similarity	99.2%	Pred. No. 2.5e-86		
Matches 708	Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	462	ATTGACATATAGCGAAGCGGAGGTGCTACAGAAAGAGCCCAAGCTTGGTCAGGGACAC	521	
Db	720	ATTGACATATAGCGAAGCGGAGGTGCTACAGAAAGAGCCCAAGCTTGGTCAGGGACAC	661	
QY	522	GGCGACTATCCTTATTTTGAACAAGTGCACAAAGATGCCACAAATGTGCAGCAGAGCTTT	581	
Db	660	GGCGACTATCCTTATTTTGAACAAGTGCACAAAGATGCCACAAATGTGCAGCAGAGCTTT	601	
QY	582	GAGGAAGCGGTTTCGAAGAGTCTTGTCTACCGAGGATAGTTCAGATCATTTGATTTCAGACA	641	
Db	600	GAGGAAGCGGTTTCGAAGAGTCTTGTCTACCGAGGATAGTTCAGATCATTTGATTTCAGACA	541	
QY	642	GACACAGTCAATCTTCAACCGAAGGCCCAAGCTAGTCACTTGTCTGTTGATTGTTAGAT	701	
Db	540	GACACAGTCAATCTTCAACCGAAGGCCCAAGCTAGTCACTTGTCTGTTGATTGTTAGAT	481	

QY	702	TGTTGATGCAATCTTAACTCAACCTCAACATATACACAAATCAACATGGGATGGAGAAGA	761
Db	480	TGTTGATGCAATCTTAACTCAACCTCAACATATACACAAATCAACATGGGATGGAGAAGA	421
QY	762	GAATTAGCGTTTGACAGCAGTGTATCACTACTAATAAATTAACCTAATGTCGTGCTTC	821
Db	420	GAATTAGCGTTTGACAGCAGTGTATCACTACTAATAAATTAACCTAATGTCGTGCTTC	361
QY	822	ATTAGTTGGTGGGAGAAAGGACACATCCACTCTTGGAGGAATATATTACTCAATAATGG	881
Db	360	ATTAGTTGGTGGGAGAAAGGACACATCCACTCTTGGAGGAATATATTACTCAATAATGG	301
QY	882	CACCTTACATTTAATAATCTGACAGTGTCTAATAACGTTCTTTTAATTTAAATATGTA	941
Db	300	CACCTTACATTTAATAATCTGACAGTGTCTAATAACGTTCTTTTAATTTAAATATGTA	241
QY	942	AGTTGCAGAGCTAATTAATGAAATGACCAAGACTTTAATTAATAAAAAATAAGAACTT	1001
Db	240	AGTTGCAGAGCTAATTAATGAAATGACCAAGACTTTAATTAATAAAAAATAAGAACTT	181
QY	1002	GACTATCTAGAGTTATACCTGGATTTTTCTGGGAAAATGGAGAACTACTCTTTTATA	1061
Db	180	GACTATCTAGAGTTATACCTGGATTTTTCTGGGAAAATGGAGAACTACTCTTTTATA	121
QY	1062	TGTGTATGTTTTTATGCAATTAGCAATTGATTTCTTGGTTCCAGGGAATACTTTCCTAAAG	1121
Db	120	TGTGTATGTTTTTATGCAATTAGCAATTGATTTCTTGGTTCCAGGGAATACTTTCCTAAAG	61
QY	1122	CAATAATGTTAGATATTAAGAATAAAATCTAATGTAAAAAATAAAAAAATAAAAAA	1175
Db	60	CAATAATGTTAGATATTAAGAATAAAATCTAATGTATTTGCAAAAAAATAAAAAA	7

RESULT 12	CD515129	881 bp	mrna	linear	EST 06-JUN-2003
LOCUS	CD515129				
DEFINITION	AGENCOURT_14365496 NIH_MGC_181 Homo sapiens cDNA clone				
	IMAGE:30396118 5', mrna sequence.				
ACCESSION	CD515129				
VERSION	CD515129.1	GI:31446847			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 881)				
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished				
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgapbs-r@mail.nih.gov Tissue Procurement: Dr. Michael Brownstein cDNA Library Preparation: Invitrogen Corp cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: http://image.llnl.gov Plate: NDAM482 row: 1 column: 07 High quality sequence stop: 714.				

```

FEATURES
source
high quality sequence stop: /14.
location/Qualifiers
1..881
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30398118"
/tissue_type="White Matter"
/dev_stage="Unknown"
/lab_host="DH103-Tcn A ( T1 and T5 phage resistances)"
/clone_lib="NIH MGC 181"

```

```

/note="Vector: pCMV-SPORT6.1; Site.1: NotI; Site.2: EcoRV
(destroyed); Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.42 kb. Library was constructed by
(invitrogen). Note: this is a NIH_MGC Library."
238 a 187 c 227 q 229 t
BASE COUNT

```

Query Match	59.9%	Score 704	DB 14	Length 881
Best Local Similarity	99.7%	Pred. No. 2.5e-86		
Matches 716	Conservative 0	Mismatches 0	Indels 2	Gaps 17
QY	58	GAGGGTCTTGAAGCTTTTGAGATTAAACATGCGCAGGAAAATCATCCTTTTAAAGTAA	117	
Db	146	GAGGGTCTTGAAGCTTTTGAGATTAAACATGCGCAGGAAAATCATCCTTTTAAAGTAA	205	
QY	118	TTCTCCTTGGAGATGGTGGAGTGTGGGAAGAGTTCACTATGAACAGATGATGTAACATA	177	
Db	206	TTCTCCTTGGAGATGGTGGAGTGTGGGAAGAGTTCACTATGAACAGATGATGTAACATA	265	
QY	178	AGTTTGATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTTAAATAAAGATTGGAG	233	
Db	266	AGTTTGATACCCAGCTCTTCCATACAAATAGGTGTGGAATTTTTAAATAAAGATTGGAG	325	
QY	238	TGGATGACATTTTGTGTACCATCCAGATTGGGACACGCGAGGTCAGAGCGATTCCGAA	297	
Db	326	TGGATGACATTTTGTGTACCATCCAGATTGGGACACGCGAGGTCAGAGCGATTCCGAA	385	
QY	298	GCCTGAGGACACCAATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTAGTTCGATG	357	
Db	386	GCCTGAGGACACCAATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTTAGTTCGATG	445	
QY	358	ATTACAAAGGTTCCAGAACTTAAAGTAACTGGAAGAAAGAAATTCATATATTATGACAGATG	417	
Db	446	ATTACAAAGGTTCCAGAACTTAAAGTAACTGGAAGAAAGAAATTCATATATTATGACAGATG	505	
QY	418	TGAAGAGCGCTGAGAGCTTTCCCTTTTGTGATCTCTGGGTAAACAAGTTGACATAAGCGAAC	477	
Db	506	TGAAGAGCGCTGAGAGCTTTCCCTTTTGTGATCTCTGGGTAAACAAGTTGACATAAGCGAAC	565	
QY	478	GGCAGGTGTCTACAGAGAAGCCCAAGCTTGGTGCAGGGACAAACGGCGACTATCCTTATT	537	
Db	566	GGCAGGTGTCTACAGAGAAGCCCAAGCTTGGTGCAGGGACAAACGGCGACTATCCTTATT	625	
QY	538	TTGAAACAGGTGCAAAAGATGCGCAAAATGTGGCAGACGCTTTGAGGAAGCGGTTCCAA	597	
Db	626	TTGAAACAGGTGCAAAAGATGCGCAAAATGTGGCAGACGCTTTGAGGAAGCGGTTCCAA	685	
QY	598	GAGTCTTGTCTACCGAGGATAGTCAGATCATTTTCGATTCACAGACACACAGTCGATCTTC	657	
Db	686	GAGTCTTGTCTACCGAGGATAGTCAGATCATTTTCGATTCACAGACACACAGTCGATCTTC	745	
QY	658	ACGAAAGCCCAAGCTAGCTCATCTTGCTGTGATTGTTAGATTGTTGATGCATTTCTAA	717	
Db	746	ACGAAAGCCCAAGCTAGCTCATCTTGCTGTGATTGTTAGATTGTTGATGCATTTCTAA	805	
QY	718	CCAACTCACACATATACACAAAATCAACATGGGAT--GGAGAAGAGAAATTAGCGTTT	773	
Db	806	CCAACTCACACATATACACAAAATCAACATGGGATGGGAGAAGAAATTGCGTTT	863	

[illegible]

Library is normalized, and was constructed by Bento Soares and M.Fatima Bonaldo."

```

BASE COUNT      214 a 164 c 134 g 261 t      5 others
ORIGIN

Query Match      58.6%; Score 688.6; DB 9; Length 778;
Best Local Similarity 96.5%; Pred. No. 3.3e-84;
Matches 722; Conservative 0; Mismatches 24; Indels 2; Gaps 2;

QY. 267 TGGACACGGCAGGTCAGGAGCGATTCGGAAGCTGAGGACACCATTTACAGAGGTTCT 326
Db 754 TTGGACACGGCAGGTTCCAGGAGCGATT-CGAAAGCTTAGGACACATTTACAGAGGTTCT 696
QY 327 GACTGCTGCTGCTTACTTTTAGTGTGATGATTCCAAAGCTCCAGACTTAAGTAAC 386
Db 695 GACTGCTGCTGCTTACTTTTAGTGTGATGATTCCAAAGCTCCAGACTTAAGTAAC 637
QY 387 TGGAAAGAAAGTAATCATATATTATGACAGATGTGAAGAGCCTGAGAGCTTTCTTTTGTG 446
Db 636 TGAAGGAAAGAAATTCATATATTATGACAGATGTGAAGAGCCTGAGAGCTTTCTTTTGTG 577
QY 447 ATTCTGGGTACAGATTGACATTAAGCGACGCGAGTGTCTACAGAGAGCCCAAGCT 506
Db 576 ATTCTGGGTACAGATTGACATTAAGCGACGCGAGTGTCTACAGAGAGCCCAAGCT 517
QY 507 TGGTGCAGGGACACACGGCGACTATCTCTATTTTGAACAAAGTGCACAAAGATGCCAAT 566
Db 516 TGGTGCAGGGACACACGGCGACTATCTCTATTTTGAACAAAGTGCACAAAGATGCCAAT 457
QY 567 GTGGCAGACGCTTTGAGGAGGCTTCGAAGATCTTCTGCTACCGAGATAGGTGAGT 626
Db 456 GTGGCAGACGCTTTGAGGAGGCTTCGAAGATCTTCTGCTACCGAGATAGGTGAGT 397
QY 627 CATTTGATTCAGACAGACACAGTCAATCTTCACCGAAAGCCCAAGCTAGCTCATCTTGC 686
Db 396 CATTTGATTCAGACAGACACAGTCAATCTTCACCGAAAGCCCAAGCTAGCTCATCTTGC 337
QY 687 TGTGATGTTAGATTGTTGATGCTTAACTCAACATCTCACATATACACAAATCAACA 746
Db 336 TGTGATGTTAGATTGTTGATGCTTAACTCAACATCTCACATATACACAAATCAACA 277
QY 747 TGGGGATGGAAGAGAAATAGAGTTTGCAGAGCTGATCATCTACTATAAATTAAC 806
Db 276 TGGGGATGGAAGAGAAATAGAGTTTGCAGAGCTGATCATCTACTATAAATTAAC 217
QY 807 TAATGTTGCTGCTTCAATAGTTGGTGGAGAGGGACACATCCACTCTTTGGAGGAATA 866
Db 216 TAATGTTGCTGCTTCAATAGTTGGTGGAGAGGGACACATCCACTCTTTGGAGGAATA 157
QY 867 TTTTACTCAATAAGGACCTTACATTTATTAATTTGAACAGTTGCTTAATAGCTTCTT 926
Db 156 TTTTACTCAATAAGGACCTTCAATTTATTAATTTGAACAGTTGCTTAATAGCTTCTT 97
QY 927 TAATTTAAATATGTAAGTTGACAGCTAATAAATGAATGACCAAGCTTTAATTAAT 986
Db 96 TAATTTAAATATGTAAGTTGACAGCTAATAAATGAATGACCAAGCTTTAATTAAT 37
QY 987 ARAAATAAGAACTTCACTATTCTAGAA 1014
Db 36 ARAAATAAGAACTTCACTATTCTAATAA 9

```

RESULT 15
 AI758228/c
 LOCUS
 DEFINITION
 Cy6b09.x1 NCI CGAP Kid11 Homo sapiens cDNA clone IMAGE:2284025 3', similar to SW:RAB9_HUMAN P51151 RAS-RELATED PROTEIN RAB-9. /, mRNA sequence.
 ACCESSION
 VERSION
 AI758228
 AI758228.1 GI:5151951
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 708)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
 Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
 Seq primer: -40UP from Gibco
 High quality sequence stop: 450.

FEATURES

source
 1..708
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clones="IMAGE:2284025"
 /lab_host="DH10B"
 /clone_lib="NCI-CGAP Kid11"
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones IDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 202 a 144 c 125 g 234 t 3 others
 ORIGIN

Query Match 58.4%; Score 686; DB 9; Length 708;
 Best Local Similarity 98.7%; Pred. No. 7.8e-84;
 Matches 700; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

```

QY 300 CTGAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTAGTGTGATGAT 359
Db 708 CTTAGGACACCATTTTACAGAGGTTCTGACTGCTGCTGCTTACTTTAGTGTGATGAT 650
QY 360 TCACAAAGCTTCAGAACTTAAGTAACCTGGAAGAAAGATTCATATATTATGAGATGTG 419
Db 649 TCACAAAGCTTCAGAACTTAAGTAACCTGGAAGAAAGATTCATATATTATGAGATGTG 590
QY 420 AAAGAGCTGAGAGCTTTCTTTTGTGATTTCTGGTTAACAGATTGCATATAGCGACGG 479
Db 589 AAAGAGCTGAGAGCTTTCTTTTGTGATTTCTGGTTAACAGATTGCATATAGCGACGG 530
QY 480 CAGTGTCTACAGAAAGCCCAAGCTTGTGTGAGGACAAACGGCGACTATCTTATTTT 539
Db 529 CAGTGTCTACAGAAAGCCCAAGCTTGTGTGAGGACAAACGGCGACTATCTTATTTT 470
QY 540 GAAACAAAGTGCACAAAGATGCCACAAATGTGGCAGAGCCCTTTGAGGAGCGGTTGAGAA 599
Db 469 GAAACAAAGTGCACAAAGATGCCACAAATGTGGCAGAGCCCTTTGAGGAGCGGTTGAGAA 410
QY 600 GTTCTTCTACCGAGGATAGGTGAGATCAATTTGATTCAGACACACAGTCAATCTTAC 659
Db 409 GTTCTTCTACCGAGGATAGGTGAGATCAATTTGATTCAGACACACAGTCAATCTTAC 350
QY 660 CGAAAGCCGAGGCTAGCTCATCTTCTGTTGATTTGTTAGATTGTTGATTCATTAACC 719
Db 349 CGAAAGCCGAGGCTAGCTCATCTTCTGTTGATTTGTTAGATTGTTGATTCATTAACC 290
QY 720 AACTCACATATACACAAAATCAACATGGGATGGAGAGAAATTAGCGTTTGACGCA 779

```

